THEORETICAL AND EMPIRICAL STUDIES ON POPULATION DYNAMICS, SPECIES INTERACTIONS AND EXTINCTIONS

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Gian Marco Palamara

aus

Italien

Promotionskomitee

Prof. Dr. Owen Petchey (Leitung der Dissertation und Vorsitz)

Prof. Dr. Lukas Keller

Prof. Dr. Gustav Delius

Dr. Matthew Smith

Zürich, 2015

A Beatrice

What pattern connects the crab to the lobster?
And the orchid to the primrose?
And all the four of them to me? And me to you?
And all the six of us to the amoeba in one direction
And to the back-ward schizophrenic in another?
Gregory Bateson

Contents

General Introduction

Ecology as a science is intrinsically quantitative and seek to understand the complex interactions between the biotic and abiotic components of the biosphere (Begon, Townsend & Harper 1986). Ecologists always recognized the need for an interdisciplinary approach that combines mathematical and biological disciplines (Codling & Dumbrell 2012). However, despite recent efforts in developing robust ecological theories (Marquet *et al.* 2014), the predictions that can be gathered from ecological models are not comparable to the predictions that are granted in other disciplines such as physics and engineering. Theoretical physics, as a major beneficiary of reductionism, has developed a powerful set of tools to describe emergent properties of complex systems. This approach consists in the description of natural systems by reducing them to the interaction of their parts. As an example, in statistical mechanics the laws of thermodynamics arise from the explanation of macroscopic properties in terms of the interaction of microscopic particles (Huang 1987). A similar approach can be adopted in ecology, and complex multispecies communities can be described by reducing them to the interaction of their parts. Species are the building blocks of ecosystems, and they form more or less complex networks of interactions (Begon, Townsend & Harper 1986). In this thesis, we aim to understand the effects of such interactions on the dynamics of populations.

We investigate the role of species' interactions (e.g. competition and predation), chance and environmental variation on population dynamics. We take the reductionist approach focusing on single species models. We will systematically use stochastic models of population dynamics in the form of birth

and death processes (BDP) (Gardiner 1985) to describe single species population dynamics. A BDP is given by two mathematical functions, the birth and the death rate of the species, which can be related to the probability of observing a given number of individuals of the species (Gardiner 1985). We include interspecific interactions in the BDPs by approximating their effect on the population dynamics of a single species. Specifically we investigate the stochastic logistic process (Nåsell 2001) varying different aspects of its formulation in order to include interspecific interactions. Interaction terms can be added on both birth and death rates of single species BDPs by assuming the density of other species (e.g. predators or competitors) is constant. We systematically investigate the range of validity of such approximations using both analytical and numerical tools. Moreover, for some of the models studied, we develop inference frameworks to obtain relevant parameters and to make reliable predictions.

Using the theory of birth and death processes we answer basic ecological questions related to environmental dependencies, invasion, and extinction dynamics. We describe different applications of the stochastic logistic process to describe how to infer the temperature dependence of ecological parameters in single species models. We describe how to predict the establishment probability of an invader species when interspecific competition with a resident species is present. We show how to compute the mean time to extinction of a species when predation is present, and we describe the combined effects of both competition, predation and environmental variability on the extinction dynamics observed in a small microbial food web.

In the last two chapters of the thesis, we focus on extinction dynamics in both single species and multispecies models of population dynamics. In fact, research shows (MEA 2005) that the present rate of extinction in all ecosystems is about 100 times higher than the extinction rate observed in the fossil record between major mass extinction events (Raup & Sepkoski 1984; Rhode & Muller 2005). Such estimates have led to forewarning of a mass extinction event. For this reason, 2010 was declared by the United Nations as the In-

ternational Year of Biodiversity. Attention of the international community towards biodiversity is motivated by an attempt at understanding the causes and the consequences of the biodiversity crisis.

In the general introduction, we give an overview of the modeling approaches used to describe both deterministic and stochastic population dynamics focusing on how extinctions have been described. We provide a theoretical framework that will be used to include species interactions in single species models. We then focus on inference methods that can be developed in this framework and on possible empirical systems that can be used to test models' predictions.

0.1 Deterministic population dynamics

The first studies of population dynamics are at least as old as the findings of Fibonacci in the XIII century for the increase of rabbit populations (Sigler 2002). The field evolved since then until the 1920's, when Alfred Lotka and Vito Volterra proposed independently a pair of first order nonlinear differential equations in which two species interact, one a predator, one its prey (Lotka 1920; Volterra 1926). These equations show that species' densities can have cycles and periodic behavior. Lotka Volterra equations in their original form present some unrealistic features. For example, a linear functional response is used to describe the influence of predation on the growth of the prey. This linearity results in neutral stability i.e., the model contains a limit cycle whose amplitude is determined by the initial populations (Drossel 2001). The study of more realistic expressions for the functional response was carried out during the 1960's. Rosenzweig and MacArthur (Rosenzweig & MacArthur 1963; Rosenzweig 1969) developed a graphical method to investigate the asymptotic behavior of dynamical equations. They found that limit cycles and fixed points are present when a saturation value is imposed for a large number of prey (The so called type-II functional response). A specific type II functional response, suggested by Holling, is largely used in

modeling (Holling 1965) because of its simple derivation from the assumption of random encounters between predators and prey. The idea was extended and new parameters describing predation, such as attack rate and handling time, were used to characterize different functional responses (Hassel 1978; Laska & Wootton 1998). In the third chapter of the thesis, we will systematically investigate the effects of the foraging parameters on the mean time to extinction of a prey species subjected to constant predation.

The generalization of the Lotka-Volterra equations with realistic functional responses to systems with more than two species was easy for food chains (Post *et al.* 2000), but not so obvious for food webs. The first to make a systematic analysis of the dynamical stability of ecological systems was May in the 1970's. In a set of papers (May 1972, 1973, 1977) he applied the classical linear stability analysis to a random model of food webs using Lotka-Volterra equations, finding what is now called the May criterion of stability. The criterion states that near an equilibrium point, food web stability scales with complexity (measured as the product of the number of species and the connectance of the food web) with exponent $-\frac{1}{2}$. In the same period, the development of chaos theory suggested limits to the predictive understanding of dynamical systems based on the time evolution of nonlinear deterministic differential equations like the Lotka-Volterra ones (May 1976). Smale and Hirsch proved that competitive Lotka-Volterra systems can generate chaos for more than four species and that in general they can exhibit any asymptotic behavior, including fixed points, limit-cycles and strange attractors(Smale 1976; Hirsch 1998). Deterministic chaotic population dynamics is characterized by positive Lyapunov exponents and, according to the time scale used to describe population models, predictability will always be limited by a time horizon. However, experimental demonstrations of chaos in ecology have been limited to laboratory systems with artificial species combinations (e.g. Benincá *et al.* (2008)).

After May's work, despite their unrealistic features (May 1974), Lotka-Volterra equations have been widely used to describe population dynam-

ics (e.g. (Cohen *et al.* 1990)). On the other hand, the desire of greater realism has resulted in several versions of the multispecies equations that go beyond the Lotka-Volterra scheme. The predation equations can be improved (Williams & Matinez 2004) incorporating parameters derived from the metabolic theory such as allometric scaling constants (Brown *et al.* 2004). Various bioenergetics equations have been proposed and coupled to structural models of food webs (Yodzis & Innes 1992; Brose, Williams & Martinez 2006; Williams 2008). Measures of permanence (Hofbauer & Sigmund 1988) and persistence (Martinez, Williams & Dunne 2005) can be performed looking at the orbit of the equations and the stability of food webs can be checked using different structural models, functional responses and dynamical parameters (Stouffer & Bascompte 2010). Findings on this field show that dynamical equations based on allometric scaling relations stabilizes the dynamics of food webs (Kartascheff *et al.* 2009; Otto, Rall & Brose 2007). We find confirmation of these findings in the third chapter of the thesis, where we show how different functional responses produce different stability patterns, even in a simple single species model.

0.2 Stochastic population dynamics

The equations described in the last section for population dynamics are deterministic. From a mathematical point of view this means that, once the initial condition is known, e.g. the initial abundances are given, the evolution of species' densities in time is unique. Species in real ecosystems, however, are subjected to demographic and environmental randomness that can lead the system to different population dynamics. Extinction risk (or invasion risk) is then influenced by numerous stochastic processes (e.g. demographic and environmental processes (Lande 1993)) leading to fluctuations in population size (De Roos & Persson 2002; Melbourne & Hastings 2008; Hakoyama *et al.* 2000). This stochastic behavior is not predictable using deterministic differential equations like the Lotka-Volterra ones (Black & McKane 2012).

The easiest way to include randomness in models of population dynamics is to add 'manually' a fluctuating term to the Lotka-Volterra equations (Leigh 1968). There are several stochastic versions of the deterministic equations which are widely used in dynamical systems' literature (Klebaner & Liptser 2001). These methods for including stochasticity present theoretical limits; in particular they don't incorporate the intrinsic stochastic nature of the process *ab initio*. As a consequence, the predictions of these models are based on the specific nature of the probability distribution used to describe fluctuations in population densities. On the other hand, stochastic models of population dynamics such as BDPs are based on *a priori* probabilistic assumptions that intrinsically include this randomness. Moreover, the deterministic equations can always be derived from the stochastic ones as a first approximation. Archetypal models that we will use in this thesis include the Verhulst logistic model (Nåsell 2001; Newman 2004) and are based on single step birth death processes. The fundamental assumption (which is the intrinsic stochastic nature of the process) is that the process has the Markov Property i.e., the conditional probability distribution of future states of the process, given the present state and the past state, depends upon only the present state. Once the Markov property holds, and birth and death rates of species are given, it is possible to obtain a master equation (ME) for the time evolution of the probability density for population sizes in a given ecosystem (Gardiner 1985). The mean time to extinction (MTE) of a given species can then be derived looking at the probability of having no individuals (Ludwig 1999; Ovaskainen & Meerson 2010). Similarly, invasion risk can be determined by looking at the probability of reaching a threshold population density (Black & McKane 2012).

When the number of individuals is sufficiently large, the system fluctuates around peaks of its probability distribution and, in first approximation, can be described by the corresponding deterministic rate equation (Ebenman, Law & Borrvall 2004; McKane, & Newman 2004; Black & McKane 2012). However, the stochastic nature of the model allows the system to experience huge fluctuations of the order of the mean population size. These rare

events, which are absent in the deterministic description, can lead to irreversible consequences such as species' extinction. Standard approximations of these stochastic models, such as the diffusion approximation (e.g. Ross, Taimre & Pollett (2006); Ross, Pagendam & Pollett (2009)), are capable of predicting the fluctuating behavior of the system. Specifically, Diffusion approximations are useful to describe systems with a large size (Gardiner 1985) that in terms of populations correspond to ecosystems with a large number of individuals for each species. We will use the diffusion approximations to obtain the likelihood function of a stochastic model in the first two chapters of the thesis.

Stochastic population dynamics is at the foundation of research concerning the factors that determine when a species is most likely to suffer extinction or to establish in a new habitat. For example, the mean time to extinction (MTE) of a single species population subjected to only demographic stochasticity grows exponentially with its carrying capacity (K) (Lande 1993). Environmental stochasticity, on the other hand, gives rise to a power law relationship between MTE and K (Lande 1993; Foley 1994). Present research about the risk of extinction has mostly considered species as isolated entities (Purvis *et al.* 2000) influenced by the abiotic environment. Other effects such as stochasticity, genetic and environmental change have been applied to this species-centric approach leading to population viability analysis (PVA) (Brooks *et al.* 2000) and to the quantification of extinction risk (Mace *et al.* 2008). PVA has sought to understand how the characteristics of individual species influence their probability of extinction and combines the effects of these different factors to estimate the overall probability that a population will go extinct (Beissinger & McCullough 2002; Mace *et al.* 2008). However, the single species models used for assessing population viability often lack the explicit incorporation of interspecific interactions (Sabo & Gerber 2007; Sabo 2008). In the second and third chapter of the thesis we will show how to incorporate interspecific interactions in single species models.

0.3 Inference methods

Increasing availability of empirical data and increasing computational power are improving our predictive understanding of ecological systems. Together with models describing deterministic and stochastic population dynamics, inference techniques to infer population parameters have been developed. In this subsection, we review some basic concepts regarding statistical inference for population time series data.

The likelihood function i.e., the probability of the data given the model, is a fundamental mathematical tool used in statistical inference. Once a proper likelihood function is derived for a population model describing empirical data (often in the form of time series data) there are two ways of obtaining the population parameters, the frequentist and the Bayesian approach (Hartig*et al.* 2011). The frequentist approach consists in Maximum Likelihood Estimation (MLE) i.e., obtaining the population parameters by directly maximizing the likelihood function (Fisher 1922). The Bayesian approach uses Bayes's rule (Price *et al.* 2009) to update the likelihood function as additional evidence is acquired. The Bayesian framework is more useful when multiple data sources can be used to infer the same population parameters and is implemented using Markov chain Monte Carlo (MCMC) methods. MCMC methods are more computationally demanding than MLE, however MCMC methods give a more complete estimation of the probability distribution of the population parameters and of their correlation.

Another important choice related to statistical inference is whether to include measurement errors. In fact, any measure of population density (or number of individuals) is affected by measurement errors such as sampling error or observational errors. A whole body of theory, named in the ecological literature as theory of "state space models", deals with uncertainties related to sampling errors in population time series data (De Valpine & Hastings 2002). State space models in both discrete and continuous time have been used to infer parameters in both laboratory and field studies (e.g. Dennis *et al.* (2006)

and Dennis & Ponciano (2014)). State space models, when applied to time series data, are based on the assumption that the true population sizes are parameters that can be inferred from the measured population sizes (Wang 2007). Therefore, state space models can be computationally demanding as they require fitting an extra parameter (the latent population state) for every data point. When the likelihood function is obtained analytically, as, for example, for single species models, the computational effort needed to run state space models can be reduced. We will use state space models to infer population parameters in the first two chapters of the thesis.

More sophisticated techniques have also been developed when the likelihood function of a population model cannot be obtained analytically (Marjoram *et al.* 2003; Hartig*et al.* 2011), as, for example, for multispecies population models. Extending stochastic models to systems with more than one species is numerically feasible but the mechanistic understanding of such more complex can be limited by their mathematical intractability. The Bayesian framework can be still used when it is not possible to obtain analytical expressions for population probability distributions. Recently several numerical techniques such as particle filters (Ionides 2003; Ionides, Bretó & King 2006) or approximate Bayesian computation (ABC) (Beaumont 2010) have been developed. Those methods simulate directly the likelihood function using the Gillespie algorithm, a well know discrete event simulator (Gillespie 1976, 1977). The likelihood of the model is then simulated at each iteration of the Markov chain (Hartig*et al.* 2011), making the corresponding inference framework more computationally demanding.

0.4 Empirical systems

When modeling real systems there is always a hierarchy of approaches: from simplest models describing well single physical phenomena that however lack of realism, to more complicated models, with more parameters, more realistic but difficult to understand (Royle & Dorazio 2008). For this reason, the same

hierarchical perspective has to be adopted when comparing models' properties to real ecological systems. Real ecosystem, at the landscape level, present a degree of complexity that cannot be incorporated in a single model as the ones presented in this thesis. Observational studies on extinctions are mostly presented in the context of the theory of island biogeography (MacArthur & Wilson 1967). These studies are based on systems with fixed boundaries such as islands and lakes (Peltonen & Hanski 1991; Burkey 1995). Extinction rates are described in function of different systems' properties such as island's area or species composition (Lawton & May 1995). Moreover, even if studying extinctions on real systems is possible but not so practical, it is impossible to use them to test models of extinctions and ecosystem collapse. In fact, in order to have a coherent statistical description of the extinction processes, species loss has to be simulated on different replicas of the same system, and this is not achievable for real ecosystems. On the other hand, laboratory systems such as protists' communities can be replicated in order to test specific ecological hypothesis.

Since the pioneering work of Gause (Gause 1932), studies carried out using microcosms communities, combine advantages of field studies such as environmental variance and realistic species combinations with advantages of small size and short generation times (Srivastava *et al.* 2004). Microcosm experiments can be designed to test model properties and are as complex and biologically realistic as other natural systems (Benton *et al.* 2007). There is a whole body of literature about empirical studies of extinctions conducted under controlled laboratory conditions (Griffen & Drake 2008). As an example, we will present a set of highly replicated experimental studies of extinction dynamics in complex microcosm communities (Worsfold *et al.* 2009). Experiments were conducted with communities of 17 species at multitrophic levels varying two environmental variables i.e., temperature and nutrient concentration. These experiments show that effect of species loss on the trajectory of secondary extinctions may be influenced by several factors such as the food web structure and the environmental conditions.

0.5 Summary

- In chapter one we investigate the temperature dependence of ecological parameters. We use a stochastic logistic model parameterized with the Arrhenius equation so that activation energy drives the temperature dependence of population parameters. We perform a systematic investigation of the effect of experimental design and inference method on the estimates of activation energy. We then apply the best performing inference methods to real data for the species *Paramecium caudatum*. We find that the fraction of habitat sampled plays the most important role in determining the relative error of the estimates of activation energy. Moreover, we find that methods that simultaneously use all time series data (direct methods) and methods that estimate population parameters separately for each temperature (indirect methods) are complementary. Indirect methods provide a clearer insight into the shape of the functional form describing the temperature dependence of population parameters; direct methods enable a more accurate estimation of the parameters of such functional forms.

- In chapter two we focus on the competitive interactions between resident species and a potential invader species. Using a stochastic version of the classic competitive Lotka-Volterra equations we assess the effect of demographic stochasticity on inferring the interaction parameter and on our ability to predict the probability of an introduced species to establish. We develop a method based on a diffusion approximation for the mean and the variance of the population size of the invader, providing a modified single species model describing the dynamics of an introduced species and its interaction with the resident species. We show how having a prior knowledge of the single species demographic parameters can improve the precision of the estimates of the interaction parameter of at least one order of magnitude.

- In chapter three we study a single species birth and death process in which the death rate includes the effect of predation. Predation is included via a general nonlinear expression for the functional response of predation to prey density. We investigate the effects of attack rate and handling time on the mean time to extinction. We find that Mean time to extinction varies by orders of magnitude when altering the foraging parameters, even when the effects of these parameters on the equilibrium population size are excluded. Our results are robust to assumptions about initial conditions and variable predator abundance.

- In chapter four we present data from a highly replicated microcosm experiment where the extinction times of 17 freshwater protists species forming a small food web has been recorded. We use survival analysis to show how interspecific dependencies and environmental factors (energy and temperature), influence the species' extinction times. The experiment shows a clear interplay between species richness, environmental factors and species interactions. We find that environmental factors are the most significant in predicting the extinction times distribution of the species in the food web. We also find that the number of competitors can be used to improve the prediction of the extinction time distributions.

0.6 Zusammenfassung

- Im ersten Kapitel untersuchen wir die Temperaturabhängigkeit von ökologischen Parametern. Wir verwenden ein stochastisches logistisches Modell mit Hilfe der Arrhenius-Gleichung, damit die Aktivierungsenergie die Temperaturabhängigkeit der Populationsparameter beeinflusst. Wir führen eine systematische Untersuchung des Einflusses der Versuchsanordnung und der Inferenz-Methode auf die Schätzungen der Aktivierungsenergie durch. Wir wenden dann die besten Inferenz Metho-

den bei realen Daten der Art *Paramecium caudatum* an. Dabei finden wir, dass die Menge des entnommenen und untersuchten Lebensraums die wichtigste Rolle bei der Bestimmung des relativen Fehlers der Aktivierungsenergie spielt. Außerdem finden wir, dass sich Methoden, die gleichzeitig alle Zeitreihendaten (direkte Methoden) benützen und Methoden, die Populationsparameter für jede Temperatur (indirekte Methoden) separat verwenden, ergänzen. Indirekte Methoden bieten einen besseren Einblick in die Beschreibung der funktionalen Form (Verteilung) der Temperaturabhängigkeit von Populationsparametern; Direktverfahren ermöglichen eine genauere Schätzung der Parameter einer solchen funktionalen Form.

- Im zweiten Kapitel konzentrieren wir uns auf die kompetitive Interaktion zwischen einer gebietsansässigen Art und einer möglichen Eindringlingsart. Mit Hilfe einer stochastischen Version der klassischen Lotka-Volterra-Gleichung bestimmen wir die Auswirkung demographischer Zufälligkeiten auf den Rückschluss der Interaktionsparameter und auf unsere Fähigkeit die Wahrscheinlichkeit vorherzusagen, ob sich eine neu eingeführte Art etabliert oder nicht. Wir entwickeln eine Methode, die auf einer Diffusionsnäherung für den Mittelwert und die Varianz der Bevölkerungsgröße des Eindringlings basiert. Damit erhalten wir ein modifiziertes Modell zur Beschreibung der Dynamik einer eingeführten Art und deren Interaktion mit einer ansässigen Art. Wir zeigen, dass die vorherige Kenntnis der demographischen Parameter einer einzelnen Art, hier die des Eindringlings, die Genauigkeit der Schätzungen des Interaktionsparameters um mindestens eine Größenordnung verbessert.

- Im dritten Kapitel untersuchen wir den Prozess von Geburt und Tod von einer einzigen Art indem die Sterberate den Effekt der Prädation miteinschliesst. Die Prädation wird über einen allgemeinen nichtlinearen Ausdruck für die funktionale Reaktion der Prädation auf die Beutedichte integriert. Wir untersuchen die Auswirkungen der Angriffsrate und Behandlungszeit auf die durchschnittliche Zeit bis die Beuteart ausstirbt. Wir finden, dass die mittlere Zeit bis zum Ausster-

ben um Grössenordnungen variiert, wenn die Raten der Futtersuche ändern, auch wenn die Auswirkungen dieser Parameter auf die Gleichgewichtspopulationsgröße ausgeschlossen werden. Unsere Ergebnisse sind robust in Bezug auf die Annahmen der Anfangsbedingungen und auf variable Raubtierdichten.

- In Kapitel vier präsentieren wir die Daten von einem hoch replizierten Mikrokosmos Experiment, bei dem die Zeiten bis zum Aussterben von 17 Süßwasser Protistenarten, die ein kleines Nahrungsnetz bilden, aufgenommen werden. Wir verwenden die Überlebensanalyse um zu zeigen, wie interspezifische Abhängigkeiten und Umweltfaktoren (Energie und Temperatur) die arteigene Aussterberate beeinflussen. Das Experiment zeigt ein klares Wechselspiel zwischen Artenreichtum, Umweltfaktoren und Interaktionen zwischen den Arten. Wir finden, dass Umweltfaktoren die wichtigste Rolle bei der Vorhersage der Zeit bis zum Aussterben der verschiedenen Arten im Nahrungsnetz spielt. Wir finden auch, dass die Anzahl der Konkurrenten gebraucht werden kann, um die Verteilungen der Aussterberaten vorherzusagen.

Bibliography

Beaumont, M. A. (2010) Approximate Bayesian Computation in evolution and ecology. *Annual Review of Ecology, Evolution, and Systematics*, **41**, 379-406.

Beissinger, S. R. & McCullough, D. R. (2002) *Population viability analysis*. University of Chicago Press, Chicago.

Begon, M., Townsend, C. R. & Harper, J. L. (1986) *Ecology from individuals to ecosystems*. Blackwell publishing, Oxford.

Benincá, E., Huisman J., Heerkloss R., Johnk K. D., Van Nes E. H., Scheffer M. & Ellner, S. P. (2008) Chaos in a long-term experiment with a plankton community. *Nature*, **451**, 822-826.

Benton, T. G., Solan M., Travis, J. M. J. & Sait, S. M. (2007) Microcosm experiments can inform global ecological problems. *Trends in Ecology & Evolution*, **22**, 516-521.

Black, A. J. & McKane, A. J. (2012) Stochastic formulation of ecological models and their applications. *Trends in Ecology & Evolution*, **27**, 337-345.

Brook, B.W., O'Grady, J. J., Chapman, A.P., Burgman M. A., Akçakaya, H. R. & Frankham, R. (2000) Predictive accuracy of population viability analysis in conservation biology. *Nature*, **404**, 385-387.

Brose, U., Williams, R. J. & Martinez, N. D. (2006) Allometric scaling enhances stability in complex food webs. *Ecology Letters*, **9**, 1228-1236.

Brown, J. H., Gillooly J. F., Allen A. P., Savage, V. M. & West, G. B. (2004) Toward a metabolic theory of ecology. *Ecology*, **85** 1771-1789.

Burkey, T. V. (1995) Extinction rates in archipelagos: implications for populations in fragmented habitats. *Conservation Biology*, **9**, 527-541.

Codling, E. A. & Dumbrell, A. J. (2012) Mathematical and theoretical ecology: linking models with ecological processes. *Interface Focus*, **2**, 144-149.

Cohen, J. E., Luczak T., Newman, C. M. & Zhou, Z. M. (1990) Stochastic structure and nonlinear dynamics of food webs: qualitative stability in a Lotka-Volterra cascade model. *Proceedings of the Royal Society B: Biological Sciences*, **240**, 607-627.

De Roos, A. M. & Persson, L. (2002) Size-dependent life-history traits promote catastrophic collapses of top predators. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 12907-12912.

De Valpine, P., & Hastings, A. (2002) Fitting population models incorporating process noise and observation error. *Ecological Monographs*, **72**, 57-76.

Dennis, B. & Ponciano, J. M. (2014) Density-dependent-state space model for population-abundance with unequal time intervals. *Ecology*, **95**, 2069-2076.

Dennis, B., Ponciano, J. M., Lele, S. R., Taper, M. L. & Staples, D. F. (2006) Estimating density dependence, process noise, and observation error. *Ecological Monographs*, **76**, 323-341.

Drossel, B. (2001) Biological evolution and statistical physic. *ArXiv*.

Ebenman, B., Law, R. & Borrvall, C. (2004) Community viability analysis: the response of ecological communities to species loss. *Ecology*, **85**, 2591-2600.

Fisher, R. A. (1922) On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **222**, 309-368.

Foley, P. (1994) Predicting extinction times from environmental stochasticity and carrying capacity. *Conservation Biology*, **8**, 124-137.

Gardiner, C. W. (1985) *Handbook of stochastic methods for physics, chemistry and natural sciences*. Springer-Verlag, Berlin.

Gause, G. F. (1932) Experimental studies on the struggle for existence. *Journal of Experimental Biology*, **9**, 389-402.

Gillespie, D. T. (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, **22**, 403-434.

Gillespie, D. T. (1977) Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, **81**, 2340-2361.

Griffen, B. D. & Drake, J. M. (2008) A review of extinction in experimental populations. *Journal of Animal Ecology*, **77**, 1274-1287.

Hakoyama, H., Iwasa, Y. & Nakanishi, J. (2000) Comparing risk factors for population extinction. *Journal of Theoretical Biology*, **204**, 327-336.

Hartig, F., Calabrese, J. M., Reineking, B., Wiegand, T. & Huth, A. (2011) Statistical inference for stochastic simulation models - theory and application. *Ecology Letters*, **14**, 816-827.

Hassell, M. P. (1978) *The dynamics of arthropod predator-prey systems*. Princeton University Press, Princeton.

Hirsch, M. (1988) Systems of differential equations which are competitive or cooperative: competing species. *Nonlinearity*, **1**, 51-71.

Hofbauer, J. & Sigmund, K. (1988) *Theory of evolution and dynamical systems*. Cambridge University Press, Cambridge.

Holling, C. S. (1965) The functional response of predators to prey density and its role in mimicry and population regulation. *Memoirs of the Entomological Society of Canada*, **45**, 1-60.

Huang, K. (1987) *Thermodynamics and statistical mechanics* Wiley.

Ionides, E. L. (2003) Inference and filtering for partially observed diffusion processes via sequential Monte Carlo. *Department of Statistics technical report*, The University of Michigan, Ann Arbor.

Ionides, E. L., Bretó, C., & King, A. A. (2006) Inference for nonlinear dynamical systems. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 18438-18443.

Kartascheff, B., Heckmann L., Drossel, B. & Guill, C. (2009) Why allometric scaling enhances stability in food web model. *Theoretical Ecology*, **3**, 195-208.

Klebaner, F. C. & Liptser, R. (2001) Asymptotic analysis and extinction in a stochastic Lotka-Volterra model. *The Annals of Applied Probability*, **11**, 1263-1291.

Landé, R. (1993) Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *The American Naturalist*, **142**, 911-927.

Laska, M. S. & Wootton, J. T. (1998) Theoretical concepts and empirical approaches to measuring interaction strength. *Ecology*, **79**, 461-476.

Lawton, J. H. & May, R. M. (1995) *Extinction rates*. Oxford University Press, Oxford.

Leigh, E. R. (1968) The ecological role of Volterra's equations. in *Some Mathematical Problems in Biology*.

Lotka, A. J. (1920) Analytical note on certain rhythmic relations in organic systems. *Proceedings of the National Academy of Sciences of the United States of America*, **6**, 410-415.

Ludwig, D. (1999) Is it meaningful to estimate a probability of extinction. *Ecology*, **80**, 298-310.

MacArthur, R. H. & Wilson, E. O. (1967) *The theory of island biogeography.* Princteon University Press, Princeton.

Mace, G. M., Collar N. J., Gaston K. J., Hilton-Taylor C., Akçakaya, H. R., Leader-Williams, N., Milner-Gulland, E. J. & Stuart, S. N. (2008) Quantification of extinction risk: IUCN's system for classifying threatened species. *Conservation Biology*, **22**, 1424-1442.

Marjoram, P., Molitor, J., Plagnol, V. & Tavaré, S. (2003) Markov chain Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 15324-15328.

Marquet, P. A., Allen, A. P., Brown, J. H., Dunne, J. A., Enquist, B. J., Gillooly, J. F., Gowaty, P. A., Green, J. L., Harte, J., Hubbell, S. P., O'Dwyer, J., Okie, J. G., Ostling, A., Ritchie, M., Storch, D. & West, G. B. (2014) On theory in ecology. *BioScience*, **64**, 701-710.

Martinez, N. D., Williams, R. J. & Dunne, J. A. (2005) Diversity, complexity, and persistence in large model ecosystem. Santa Fé Institute, 163-185.

Millennium Ecosystem Assessment, (2005) *Ecosystem and human well being.* World Resources Institute.

May, R. M. (1972) Will a large complex system be stable? *Nature*, **238**, 413-414.

May, R. M. (1973) Qualitative stability in model ecosystems. *Ecology*, **54**, 638-641.

May, R. M. (1974) *Stability and complexity in model ecosystems.* Princeton Landmarks in Biology, Princeton.

May, R. M. (1976) Simple mathematical models with very complicated dynamics. *Nature*, **261**, 459-467.

May, R. M. (1977) Thresholds and breakpoints in ecosystems with a multiplicity of stable states. *Nature*, **269**, 471-477.

McKane, A. J. & Newman, T. J. (2004) Stochastic models in population biology and their deterministic analogs. *Physical Review E*, **70**, 1-19.

Melbourne, B. A. & Hastings, A. (2008) Extinction risk depends strongly on factors contributing to stochasticity. *Nature*, **454**, 100-103.

Nåsell, I. (2001) Extinction and quasi-stationarity in the Verhulst logistic model. *Journal of Theoretical Biology*, **211**, 11-27.

Newman, T. J., Ferdy, J. B. & Quince, C. (2004) Extinction times and moment closure in the stochastic logistic process. *Theoretical Population Biology*, **65**, 115-126.

Otto, S. B., Rall, B. C. & Brose, U. (2007) Allometric degree distributions facilitate food-web stability. *Nature*, **450**, 1226-1229.

Ovaskainen, O. & Meerson, B. (2010) Stochastic models of population extinction. *Trends in Ecology & Evolution*, **25**, 1-42.

Peltonen, A. & Hanski, I. (1991) Patterns of island occupancy explained by colonization and extinction rates in shrews. *Ecology*, **72**, 1698-1708.

Pimm, S. L., Lones, H. L. & Diamond, J. (1988) On the risk of extinction. *The American Naturalist*, **132**, 757-785.

Post, D. M., Conners, M. E. & Goldberg, D. S. (2000) Prey preference by a top predator and the stability of linked food chains. *Ecology*, **81**, 8-14.

Price, C. A., Ogle, K., White, E. P. & Weitz, J. S. (2009) Evaluating scaling models in biology using hierarchical Bayesian approaches. *Ecology Letters*, **12**, 641-651.

Purvis, A., Gittleman J. L., Cowlishaw, G. & Mace, G. M. (2000) Predicting extinction risk in declining species. *Proceedings of the Royal Society B: Biological Sciences*, **267**, 1947-1952.

Raup, D. M. & Sepkoski, J. J. (1984) Periodicity of extinctions in the geologic past. *Proceedings of the National Academy of Sciences of the United States of America*, **81**, 801-805.

Rohde, R. A., & Muller, R. A. (2005) Cycles in fossil diversity. *Nature*, **434**, 208-210.

Rosenzweig, M. L. & MacArthur, R. H. (1963) Graphical representation and stability conditions of predator-prey interactions. *The American Naturalist*, **97**, 209-223.

Rosenzweig, M. L. (1969) Why the prey curve has a hump. *The American Naturalist*, **103**, 81-87.

Ross, J. V., Pagendam, D. E. & Pollett, P. K. (2009) On parameter estimation in population models II: multi-dimensional processes and transient dynamics. *Theoretical Population Biology*, **75**, 123-132.

Ross, J. V., Taimre, T. & Pollett, P. K. (2006) On parameter estimation in population models. *Theoretical Population Biology*, **70**, 498-510.

Royle, J. A. & Dorazio, R. M. (2008) *Hierarchical modelling and inference in ecology*. Elsevier, Oxford.

Sabo, J. L., & Gerber, L. R. (2007) Predicting extinction risk in spite of predator-prey oscillations." Ecological Applications **17**, 1543-1554.

Sabo, J. L. (2008) Population viability and species interactions: Life outside the single-species vacuum. *Biological Conservation*, **141**, 276-286.

Sigler, L. E. (2002) *Fibonacci's Liber Abaci*. Springer-Verlag, Berlin.

Smale, S. (1976) On differential equations of species in competition. *Journal of Mathematical Biology*, **3**, 5-7.

Srivastava, D. S., Kolasa, J., Bengtsson, J., Gonzalez, A., Lawler, S. P., Miller, T. E., Munguia, P., Romanuk, T., Schneider, D. C. & Trzcinski, M. K. (2004) Are natural microcosms useful model systems for ecology. *Trends in Ecology & Evolution*, **19**, 379-384.

Stouffer, D. B. & Bascompte, J. (2010) Understanding food-web persistence from local to global scales. *Ecology Letters*, **13**, 154-161.

Volterra, V. (1926) Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. *Memorie Accademia dei Lincei, Roma*, **2**, 31-113.

Wang, G. (2007) On the latent state estimation of nonlinear population dynamics using Bayesian and non-Bayesian state-space models. *Ecological Modelling*, **200**, 521-528.

Williams, R. J. (2008) Effects of network and dynamical model structure on species persistence in large model food webs. *Theoretical Ecology*, **1**, 141-151.

Williams, R. J. & Martinez, N. D. (2004) Stabilization of chaotic and non-permanent food-web dynamics. *The European Physical Journal B*, **38**, 297-303.

Worsfold, N. T., Warren, P. H. & Petchey, O. L. (2009) Context-dependent effects of predator removal from experimental microcosm communities. *Oikos*, **118**, 1319-1326.

Yodzis, P. & Innes, S. (1992) Body size and consumer-resource dynamics. *The American Naturalist*, **139** 1151-1175.

Chapter 1

Inferring the temperature dependence of ecological parameters: the effect of experimental design and inference algorithm

Gian Marco Palamara, Dylan Z. Childs, Christopher F. .Clements, Owen L. Petchey, Marco Plebani & Matthew J. Smith

Ecology & Evolution, In Press.

1.1 Abstract

Understanding and quantifying the temperature dependence of population parameters, such as intrinsic growth rate and carrying capacity, is critical for predicting the ecological responses to environmental change. Many studies provide empirical estimates of such temperature dependencies, but a thorough investigation of the methods used to infer them has not been performed yet. We created artificial population time series using a stochastic logistic model parameterized with the Arrhenius equation so that activation energy drives the temperature dependence of population parameters. We simulated different experimental designs and used different inference methods, varying the likelihood functions and other aspects of the parameter estimation methods. Finally, we applied the best performing inference methods to real data for the species *Paramecium caudatum*.

The relative error of the estimates of activation energy varied between 5 and 30 %. The fraction of habitat sampled played the most important role in determining the relative error; sampling at least 1% of the habitat kept it below 50%. We found that methods that simultaneously use all time series data (direct methods) and methods that estimate population parameters separately for each temperature (indirect methods) are complementary. Indirect methods provide a clearer insight into the shape of the functional form describing the temperature dependence of population parameters; direct methods enable a more accurate estimation of the parameters of such functional forms. Using both methods, we found that growth rate and carrying capacity of *Paramecium caudatum* scale with temperature according to different activation energies.

Our study shows how careful choice of experimental design and inference methods can increase the accuracy of the inferred relationships between temperature and population parameters. The comparison of estimation methods provided here can increase the accuracy of model predictions, with important implications in understanding and predicting the effects of temperature on

the dynamics of populations.

Key Words: Activation energy, Arrhenius equation, Maximum likelihood, MCMC, Metabolic theory, Microcosm experiments, State space models, Stochastic simulations.

1.2 Introduction

Explaining the distribution and abundance of organisms requires knowledge of the environmental dependence of organismal properties (Hall, Stanford & Hauer 1992; Ives 1995), including biological rates such as birth and death rate (Volkov *et al.* 2003). Furthermore, predicting the effects of environmental change on populations benefits from understanding the environmental dependence of biological processes (Ives 1995; Thomas *et al.* 2004; Deutsch *et al.* 2008; Vasseur *et al.* 2014). Empirical relationships between the rates of physiological processes and one particularly important environmental variable, temperature, have been documented for many processes and taxa (Gillooly *et al.* 2001, 2002; Dell, Pawar & Savage 2010), including rates of food ingestion by individuals (Englund *et al.* 2011; O'Connor, Gilbert & Brown 2011; Dell, Pawar & Savage 2013), rates of population growth (Savage *et al.* 2004), and rates of various ecosystem processes (Ernest *et al.* 2003; Allen, Gillooly & Brown 2005; Yvon-Durocher *et al.* 2012). These and other relationships have been used to predict effects of temperature on population dynamics (Vasseur & McCann 2005). The overall aim of this paper is to provide improved inference methods for estimating such relationships.

Methods used to infer the population parameters from time series data typically range from classic maximum likelihood estimation (Hilborn 1997) to Bayesian inference for partially observed Markov processes (Knape & De Valpine 2012; Dennis & Ponciano 2014). When estimating population parameters, one needs a description of the sampling error associated with any experiment or field survey, as well as an explicit model of the dynamics (De Valpine & Hastings 2002; Dennis *et al.* 2006). An important decision is thus

whether inference method should explicitly account for the sampling process i.e., the process that provides the actual counts of the number of individuals. Unless the entire habitat is sampled (so that every individual is counted) the observed number of individuals will be a sample of the actual abundance (De Valpine & Hastings 2002; Dennis *et al.* 2006; Ross 2012) and not including sampling error can lead to erroneous parameter estimates (Ionides, Bretó & King 2006). Fitting stochastic population dynamic models to observed data while taking into account sampling error is a non-trivial endeavor (Ionides, Bretó & King 2006; Ross 2012). Hence, it would be very useful to know when such an approach is necessary, and when a simpler approach (e.g., a deterministic model with no accounting for sampling error) provides sufficiently accurate and precise estimates.

We focus on improving inference of the relationship between two population parameters (intrinsic growth rate r and carrying capacity K) and temperature. The Arrhenius law, which was originally proposed to describe the temperature dependence of the specific reaction rate constant in chemical reactions (Van't Hoff 1884; Arrhenius 1889), is used to describe the temperature dependence of whole-organisms metabolic rates such as growth rate (Schoolfield, Sharpe & Magnuson 1981). The Arrhenius law predicts that the natural logarithm of mass-corrected metabolic rates is a linear function of the inverse absolute temperature. The slope of this relationship gives the activation energy of metabolism (Arrhenius 1889; Schoolfield, Sharpe & Magnuson 1981) and the intercept gives the natural logarithm of the normalization constant (Brown *et al.* 2004). The temperature dependence of r has been studied extensively (Dell, Pawar & Savage 2010; Corkrey *et al.* 2012), especially in microbes (Monod 1942; Weisse & Montagnes 1998; Weisse *et al.* 2002; Price & Sowers 2004; Jang & Morin 2004; Krenek, Berendonk & Petzoldt 2011; Krenek, Petzoldt & Berendonk 2012), rotifers (Montagnes *et al.* 2001), algae (Montagnes & Franklin 2001) and insects (Irlich *et al.* 2009; Amarasekare & Sifuentes 2012). The temperature dependence of K has received less attention (Yodzis & Innes 1992; Brown *et al.* 2004; Savage *et al.* 2004; Vasseur & McCann 2005). In this study, we focus on the statistical

methods used to infer such temperature rate relationships. We do not enter the debate about the validity of Arrhenius law (Knies & Kingsolver 2010) or on the exact value of activation energy (Glazier 2006), although in the Discussion we will indicate how our insights can be used to address these debates.

Data needed to assess the temperature dependence of population parameters come in the form of time series collected at different (fixed) temperatures (Jang & Morin 2004; Beveridge, Petchy & Huphries 2010; Leary, Rip & Petchey 2012; Krenek, Petzoldt & Berendonk 2012). This is done in experiments in which single species populations are grown at a variety of temperatures, starting from very low abundances, until carrying capacity is reached. Population size is recorded with a certain temporal frequency, most often from a sub-sample of the total habitat (i.e., the population is sampled), thus providing a time series for each temperature. The estimates of r and K obtained at each temperature over a range of temperatures are used to estimate activation energy through the Arrhenius law (Gillooly *et al.* 2002; Savage *et al.* 2004). Although our study assumes a temperature range for which the Arrhenius law is appropriate, the results will generalize to a wider range of temperatures. We term the use of this approach an "indirect method" of estimating the activation energy. This is, to date, the most common approach to estimating activation energy from growth processes (Weisse *et al.* 2002; Price & Sowers 2004; Savage *et al.* 2004; Angiletta 2006; Huang, Hwang & Phillips 2011; Krenek, Berendonk & Petzoldt 2011; Corkrey *et al.* 2012; Krenek, Petzoldt & Berendonk 2012), and from other processes (Rall *et al.* 2009; Englund *et al.* 2011). An alternative approach, which we term the "direct method", is to directly fit a model of the temperature dependence of population dynamics to the entire dataset i.e., to fit to population dynamics from all the temperature treatments simultaneously. Based on limited previous comparisons of indirect and direct estimation methods, we expect the direct method to have higher accuracy and precision than the indirect method (Schoolfield, Sharpe & Magnuson 1981; Price & Sowers 2004), because it is combining more information directly in the inference process to infer fewer parameters. As well as making this comparison, we illustrate the

ecological consequences of the observed differences in accuracy and precision.

In addition to choices about inference methods, a researcher makes choices about the design of the experiments used to produce the observed data. Here we assess the importance of different experimental designs and inference methods on the ability to infer activation energy from time series data on single species experimental microcosms. We assess the performance of different inference methods given particular choices of experimental designs by estimating the activation energy of simulated population data. We also demonstrate an application of the methods to real data from experiments with *Paramecium caudatum*, a well-studied freshwater protist species (Krenek, Berendonk & Petzoldt 2011; Krenek, Petzoldt & Berendonk 2012) (Figure 1.1). We used only one species as a case study because the focus of our study is methodological, rather than descriptive. We chose *Paramecium caudatum* because it shows population growth that is well captured by the stochastic logistic equation (Leary & Petchey 2009). We provide advice for experimentalists about the most relevant factors affecting the precision and the accuracy of the estimates of activation energy for different inference methods.

To our knowledge there has been no thorough and systematic exploration of the relative importance of these issues (i.e., influence of experimental design, sampling design, model type, and inference method) for the accuracy and precision of estimates of environmental dependence of ecological parameters such as the temperature dependence of intrinsic growth rate and carrying capacity. The methods are illustrated with estimation of r and K, but can be generalized to estimation of the activation energy of other biological rates, such as maximum consumption rate (Rall *et al.* 2009; Englund *et al.* 2011), and effects of environmental variables other than temperature, for example nutrient availability (Weisse *et al.* 2002; Price & Sowers 2004).

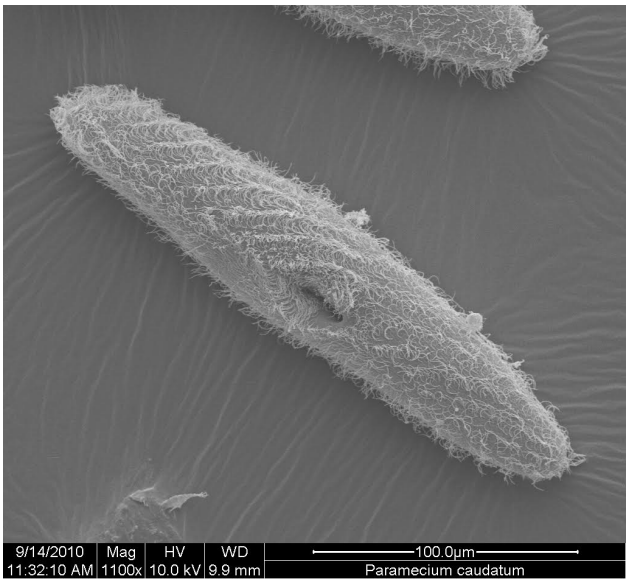

Figure 1.1: Picture of the freshwater living species *Paramecium caudatum* (courtesy of Dr. Renate Radek)

1.3 Methods

We describe population dynamics using a continuous time, stochastic logistic model (Nåsell 2001), a generalization of the deterministic logistic equation in continuous time (McKane, & Newman 2004; Gardinier 2009). Stochastic models can provide fundamentally different results from their deterministic counterparts (Ebenman, Law & Borrvall 2004; McKane, & Newman 2005), and provide a more detailed description of the mechanisms affecting population dynamics (Black & McKane 2012). For example, the carrying capacity (K) in the deterministic logistic equation is the equilibrium population density of a given species, namely the maximum sustainable population size given the available resources (Malthus 1798; Turchin 2003). Conversely, in stochastic logistic growth models K represents the mean of a long term stationary distribution around which the population fluctuates (Nåsell 2001; Dennis *et al.* 2006).

We performed a simulation study to assess the importance of experimental protocols and inference methods on the ability to estimate the activation

energy for the temperature dependence of population parameters. This involved simulating population dynamic data using a model with known activation energy in section 1.3.1 and comparison of this true activation energy to that obtained by various inference methods in section 1.3.2. We illustrated the best performing methods by estimating activation energy from real population dynamic data of a free living freshwater protist species, *Paramecium caudatum* in section 1.3.3.

1.3.1 Model and simulations

We used a simple stochastic birth and death processes (BDP) model to generate time series data of population dynamics

$$\begin{aligned} B(n;\theta) &= \theta_1(T)n\left(1-\theta_2(T)\frac{n}{N}\right), \text{and} \\ D(n;\theta) &= \theta_3(T)n, \end{aligned} \tag{1.1}$$

where $0 \leq n \leq N$ is the (integer) number of individuals, N is population size at which there is zero probability of births, θ_1 and θ_3 are the per capita birth and death rates in the absence of density dependence, respectively (units: day^{-1}), θ_2 controls the strength of density dependent effects on the probability of births (dimensionless), and (T) indicates that all θ parameters are dependent on temperature, T (measured in Kelvin). We used the BDP 1.1 because it allows to take into account all biological mechanisms affecting population dynamics (for more details on the model see 1.6.1); for simplicity we assume that density dependence only affects probabilities of births, although in reality density dependence likely influences the probability of both births and deaths (i.e. both births and deaths in process 1.1 would be influenced by N). We introduce temperature dependence to the θ parameters using the Arrhenius equation (Gillooly *et al.* 2001)

$$\theta_i(T) = \theta_{i0} \exp \frac{E_{A,i}(T-T_0)}{k_B T T_0}, \tag{1.2}$$

where $i = 1, 2, 3$ denotes the population parameter in the BDP 1.1, $E_{A,i}$ is the activation energy (units: *ElectronVolts*) for parameter θ_i, k_B is the Boltzmann constant, and T_0 is a reference baseline temperature, which we assume to be $301.15K$ ($28°C$). For most of our analyses we assume the same $E_{A,i}$ for all parameters.

The mean population abundance over time follows the logistic equation

$$\frac{dn(t)}{dt} = B(n) - D(n) = r(T)n\left(1 - \frac{n}{K(T)}\right), \tag{1.3}$$

where $r(T) = \theta_1(T) - \theta_3(T)$ is the maximum population growth rate and $K(T) = \frac{\theta_1(T)-\theta_3(T)}{\theta_1(T)\theta_2(T)}N$ is the carrying capacity (Nåsell 2001). The temperature dependencies of growth rate and carrying capacity are thus

$$r(T) = r_0 e^{\frac{E_A(T-T_0)}{k_b T T_0}}, \tag{1.4}$$

$$K(T) = K_0 N e^{\frac{-E_A(T-T_0)}{k_b T T_0}}, \tag{1.5}$$

where $r_0 = \theta_{10} - \theta_{30}$ and $K_0 = \frac{\theta_{10}-\theta_{30}}{\theta_{20}\theta_{30}}$ are the growth rate and carrying capacity at T_0. Expressions 1.4 and 1.5 indicate that growth rate and carrying capacity should increase and decrease with temperature, respectively (Savage *et al.* 2004).

We simulated the process 1.1 and the relations 1.2 using the well known Gillespie algorithm (Gillespie 1976) (see fig. 2.2 for examples). This produced continuous time series recording the exact times of individual birth and death events. To make simulated data more representative of experimental data we then sampled population size at discrete times as if only a fraction of the population had been sampled and counted (examples are shown in fig. 2.2). To simulate sampling we assumed that the numbers measured were drawn from a Poisson distribution centered on the expected number of individuals contained in a sample from the population, where the sample size FRACSAMP is the fraction of the habitat searched. We do not include an additional source of error from the imperfect ability of observers to count all

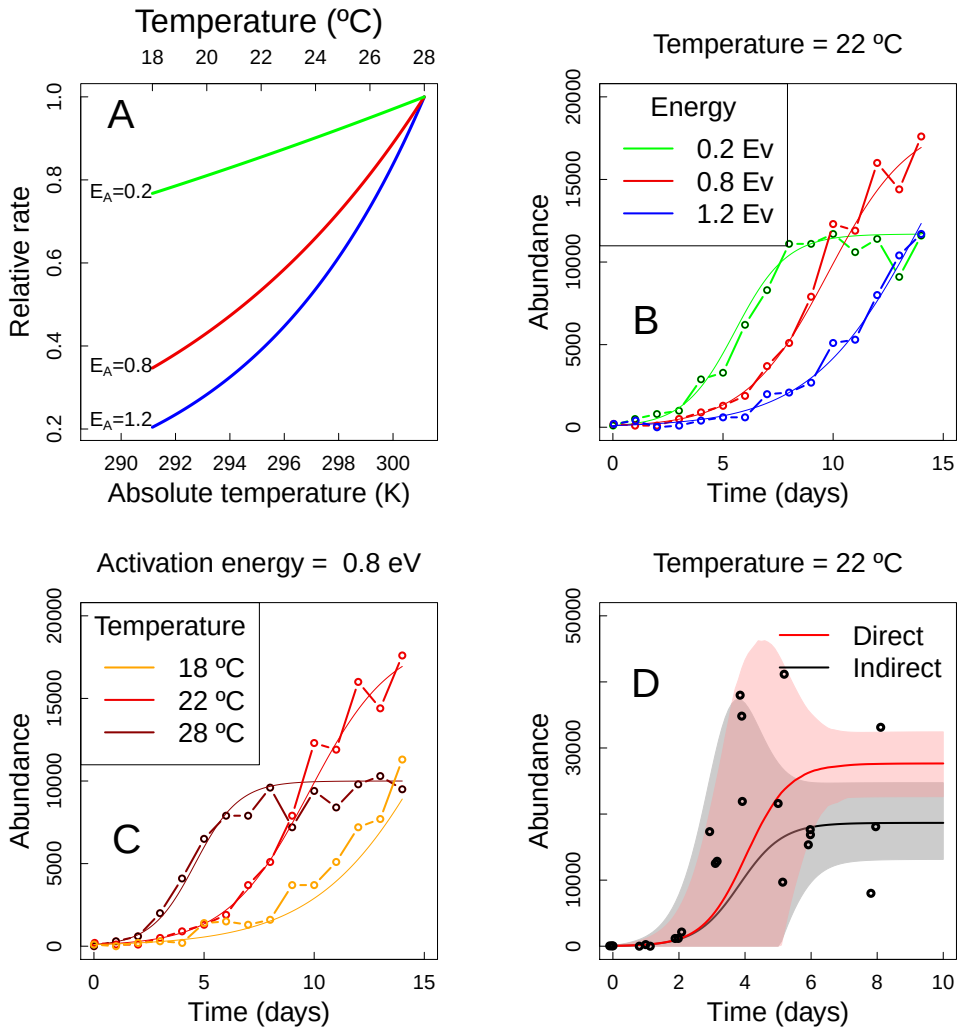

Figure 1.2: Example of temperature dependence of a rate for three different activation energies (Panel A), standardized to have the same value at $301.15K$ (Huey & Kingsolver 2011). Panels B and C show the effect of activation energy (panel B) and temperature (panel C) on time series originated by the BDP 1 with parameters scaled using equation 2. The simulated time series all have an initial condition of 100 individuals, are sampled every day for 15 days (TIMESAMP=15) and are subjected to demographic noise and sampling error (FRACSAMP=0.01). The continuous lines show the deterministic solution 13. Panel D shows real time series data (black dots) for three replicates of Paramecium caudatum monocultures (maximum FRACSAMP = 0.001). We show the corresponding fitted means (continuous lines) and modeled variances (shaded areas) using both direct (red) and indirect (black) methods. The estimated activation energies are shown in figure 1.6 and 1.8.

individuals in a sample, thus demographic stochasticity and sampling error are the only two sources of stochasticity in our simulated experimental data.

We chose parameter values for equations 1.1 and 1.2 that lead to similar simulated population dynamics to those observed in laboratory experiments (see Fig. 2.2) and that are consistent with previously published values (Savage *et al.* 2004). We set the reference temperature $T_0 = 28°C$ and scaled the other population parameters relative to their probabilities at that temperature: $\theta_1(T_0) = 1.5day^{-1}$, $\theta_2(T_0) = 1$, $\theta_3(T_0) = 0.5day^{-1}$. The population size at which the probability of births is zero, N, was fixed throughout this study to $N = 15000$ individuals. The importance of this value is detailed in the discussion, and here was chosen in order to represent a typical laboratory experiment with a microcosm of 10 ml.

These choices lead to a maximum population growth rate of $r(T_0) = 1day^{-1}$ and a minimum carrying capacity of $K(T_0) = 10000$ individuals. All simulations began with an initial population size of $n_0 = 100$ individuals and lasted 15 days. We simulated equations 1.1 and 1.2 under 81 different sets of experimental conditions, representing the range of experimental strategies likely to be considered when conducting laboratory experiments to estimate activation energy. These 81 experiments arise from a fully factorial experimental design in which four factors are varied, with three different values each. We varied

- The number of different temperatures considered, TEMPSAMP. We generated time series at 11 different temperatures from $18°C$ to $28°C$ in steps of $1°C$ but varied the numbers of different temperatures used in the estimation of activation energy: either using all 11 temperatures, using only six different temperatures (from $18°C$ to $28°C$ in steps of $2°C$), or using just three different temperatures ($18°C$, $23°C$ and $28°C$). Those temperature gradients were chosen in order to capture the temperature range where we expect the Arrhenius law 1.2 to be valid.

- The number of replicate experiments at each temperature and activa-

tion energy, REPS. We considered one, three or five replicates at each temperature. While estimation using one replicate per temperature is possible, from three to five are typically used in experiments where population time series are recorded (Leary & Petchey 2009; Krenek, Berendonk & Petzoldt 2011).

- The number of samples taken during an experiment, TIMESAMP. We considered once every three days (TIMESAMP=5), twice every three days (TIMESAMP=10) or once a day (TIMESAMP=15) over the course of each 15 days experiment. Fifteen days was sufficient to capture both the growth phase and the equilibrium phase (carrying capacity) of the population dynamics.

- The fraction of habitat sampled, FRACSAMP. We considered 1%, 0.5% and 0.1% of the entire habitat (FRACSAMP = 0.01, 0.005, 0.001), reproducing the typical search effort of experiments (De Valpine & Hastings 2002; Dennis *et al.* 2006).

For each experimental design, we then estimate activation energy using different methods.

1.3.2 Parameters inference

To conduct parameter inference we need a mathematical function defining the probability of a set of parameters given the data i.e., the likelihood function. We compared different methods for inferring activation energy (summarized in table 1.1) using five different likelihood functions (for details on the derivation of the likelihood functions see 1.6.2). The model underpinning methods M1 and M2 is the solution of equation 3.6 i.e., the likelihood function is parameterized using only the mean population abundance over time, assuming that the dynamics are deterministic. The second model (underpinning methods M3-M6) assumes that the dynamics are demographically stochastic but that there is no sampling error; the correspondent likelihood function is parameterized using both the mean and the variance of population abundance

Table 1.1: Methods to infer activation energy: Column three (parameter used) specifies which parameter is used to obtain the estimate of activation energy. Column five (Method) refers to the statistical framework used i.e., MLE (Maximum Likelihood Estimation) or MCMC (Markov Chain Monte Carlo). Column six (Corr.) states if the correction for sampling error was implemented (YES) or not (NO). The last column of the table shows computational times of each method when inferring activation energy using the same simulated data for all inference methods (FRACSAMP = 0.01, REP = 5, TIMESAMP = 10, TEMPSAMP = 11) for a fixed activation energy ($E_A = 0.2Ev$). The computational time was measured on a desktop computer whose processor is Intel(R) Xenon(R) E5645 2.4Ghz, with installed RAM of 12 GB. The numbers denoted by * are widely variable even on the same operating system. In fact, frequently the algorithm returns NA for the mean and or the variance of the parameter estimates and the time taken to obtain the parameter estimates are highly variable. The numbers reported are chosen as representative from the runs that reported real numbers for the mean and variance of the parameter estimates.

Inference Method	Likelihood Function	Parameter used	Estimate	Method	Corr.	Comp. Time
M1	$L_{phen}(\mathbf{\Theta})$ (1.15)	$log(r(T))$	INDIRECT	MLE	NO	$0.5h^*$
M2	$L_{phen}(\mathbf{\Theta})$ (1.15)	$log(K(T))$	INDIRECT	MLE	NO	$0.5h^*$
M3	$L_1(\theta')$ (4.1)	$log(r(T))$	INDIRECT	MLE	NO	$0.5h^*$
M4	$L_1(\theta')$ (4.1)	$log(K(T)/r(T))$	INDIRECT	MLE	NO	$0.5h^*$
M5	$L_1(\theta')$ (4.1)	$log(r(T))$	INDIRECT	MCMC	NO	$1h$
M6	$L_1(\theta')$ (4.1)	$log(K(T)/r(T))$	INDIRECT	MCMC	NO	$1h$
M7	$L_2(\theta')$ (1.18)	$log(r(T))$	INDIRECT	MCMC	YES	$2h$
M8	$L_2(\theta')$ (1.18)	$log(K(T)/r(T))$	INDIRECT	MCMC	YES	$2h$
M9	$L_1^D(\theta'_{\mathbf{0}}, \theta_4)$ (1.19)	$log(E_A)$	DIRECT	MCMC	NO	$1.5h$
M10	$L_2^D(\theta'_{\mathbf{0}}, \theta_4)$ (1.20)	$log(E_A)$	DIRECT	MCMC	YES	$2.5h$

(see 1.6.1 and (Ross, Pagendam & Pollett 2009) for the diffusion approximation used in the derivation of the population variance). In methods M7-M8 we add to the likelihood function of methods M3-M6 a correction taking into account for the sampling error.

Methods M1-M8 are defined as indirect as they adopt the common approach of inferring activation energy indirectly i.e., population growth rates (r) or carrying capacities (K) are inferred at different temperatures. Activation energy is then deduced from the relationship between these parameters and the inverse energy $1/k_bT$ (see figure 2.2 B) given by

$$log(K(T)/r(T)) = C_1 + 2E_A\frac{1}{k_bT}, \quad (1.6)$$

$$log(r(T)) = C_2 - E_A\frac{1}{k_bT}, \quad (1.7)$$

where $C_1 = log(N/\theta_{10}\theta_{20}) - 2E_A/k_BT_0$ and $C_2 = log(\theta_{10} - \theta_{30}) + E_A/k_BT_0$ are two temperature independent constants. Activation energy is the slope of these relationships, derived using standard linear regression between the logarithm of the parameters of the logistic equation and the inverse temperature (Schoolfield, Sharpe & Magnuson 1981) (see figure 2.2 B), as it has been extensively performed in previous studies (Schoolfield, Sharpe & Magnuson 1981; Gillooly *et al.* 2001, 2002; Savage *et al.* 2004).

The other approach we take is to infer activation energy directly. Method M9 is a generalization of methods M5-M6 and its likelihood is obtained by summing the likelihood underpinning methods M5-M6 over all observed temperatures. Similarly, method M10 is a generalization of methods M7-M8 and takes into account the sampling error. The likelihood of method M10 is obtained by summing the likelihoods of models M7-M8 over all observed temperatures (see section 1.6.2 for more details on the direct methods). The indirect methods used to infer activation energy are characterized by the choice of one parameter (growth rate or carrying capacity) whose temperature dependence (relations 1.6 and 1.7) provides an estimate of activation

energy. Direct methods, on the other hand, provide an estimate of activation energy from the global temperature dependency of all the parameters of model 1.1.

For each inference algorithm and experiment we measured the relative error (R) and precision (P) of the estimate given by

$$R = \frac{E_A - m(E_A)}{E_A}; P = \frac{se(E_A)}{E_A}, \tag{1.8}$$

where E_A is the real value of activation energy used to produce the simulated data, $m(E_A)$ is the mean of the estimate, and $se(E_A)$ is the standard error of the estimate. The accuracy of the estimates of activation energy is given by the inverse of the relative error R. When performing MLE, all the distributions of the parameters were assumed Gaussian and the standard deviation was automatically inferred, while, when performing MCMC, we always checked the shape of the distribution to be a Gaussian, especially when performing the linear regressions 1.6 and 1.7 in the indirect models. Note that an increase in precision and accuracy correspond to a decrease in the percentage given, in other words, high accuracy and precision correspond with low values of R and P.

We then applied classification and regression tree analysis (CART) (Ripley 2007) to the absolute value of the relative error of the estimates of activation energy (the response variable) for each of the methods in table 1.1, in order to assess the relative importance of different experimental factors (the explanatory variables) and their interaction (1.3). A regression tree is constructed by repeated splits of the data into mutually exclusive groups. Each split is defined by values less than some chosen value of one of the experimental factors. At each split, the data is partitioned into two groups as homogeneous as possible. Each group is distinguished by the mean of the absolute value of the relative error of the estimate of activation energy and the values of the experimental factors that define it (De'ath & Fabricius 2000; Ripley 2007). Splits are chosen in order to minimize the sum of squared error be-

tween the observation and the mean in each node of the tree. The splitting procedure is then applied to each group separately partitioning the response into homogeneous groups, and keeping the tree sensibly small. Appropriate tree size is determined setting a threshold in the reduction in homogeneity measure (De'ath & Fabricius 2000). Regression trees are a powerful tool for their capacity of interactive exploration and description of different subsets of the data and are often used instead of more classic linear model analysis (De'ath & Fabricius 2000).

1.3.3 Case study

As a case study we present data from a microcosm experiment (Leary & Petchey 2009) in which time series of abundance were collected along a gradient of six different temperatures between $18°C$ and $28°C$, where there were three replicates and TEMPSAMP=6 (please see (Leary & Petchey 2009) for supplementary detail). In this case study the fraction of habitat searched (FRACSAMP) and the frequency of sampling (TIMESAMP) were variable, the latter depending on the temperature and the former depending on the observed density; this was accounted for in the likelihood functions. We estimated the activation energy of the protist species *Paramecium caudatum* in these microcosm experiments using methods M1, M2, M7, M8 and M10 (see table 1 for definitions). Methods M7, M8 and M10 were used because we found them to be the most effective in estimating activation energy. Methods M1 and M2 (using the phenomenological likelihood 15, section 5.2) were included to act as a comparison with the best performing methods because we wanted to investigate how important their lack of accuracy and precision could be when estimating activation energy (see figure 1.6 B). We also found that real data do not strictly obey to the theory presented in (Savage *et al.* 2004) for carrying capacity (see figure 1.6 B), for this reason, while using model M10, we implemented a likelihood with two different activation energies, one for growth rate ($E_{A,r}$) and one for carrying capacity ($E_{A,K}$).

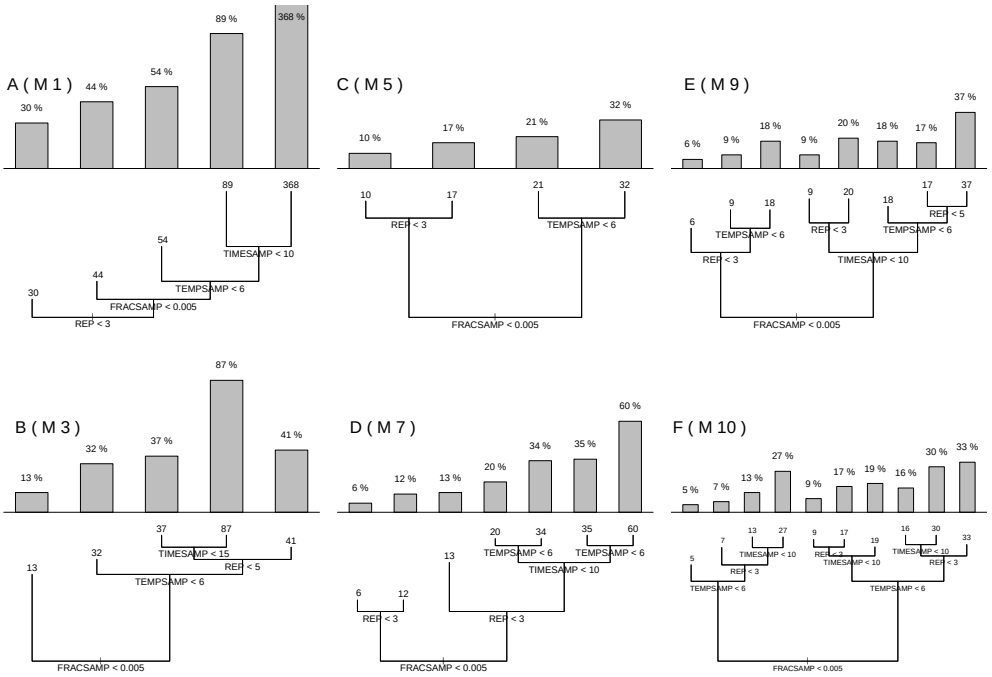

Figure 1.3: The results of the classification and regression tree (CART) analysis (Ripley 2007) of the relative error of the estimates of activation energy. The number at the leaves of the tree indicates the mean percentage value of the relative error of the estimate (see expression 8) over all the simulated experiments, following partitioning of the data in the manor specified by the tree. The threshold above each node indicates the split criterion used to separate the data. To each tree is associate a bar-chart showing the mean percentage value of each leaf. The six panels correspond to six of the models specified in table 1: model M1 (panel A), M3 (panel B), M5 (panel C), M7 (panel D), M9 (panel E), M10 (panel F).

1.4 Results

Activation energy was estimated with a wide range of accuracies across the different experimental conditions and inference methods considered, varying from high accuracy (relative error estimates being within $< 5\%$ of the mean value on average) to low accuracy (relative error estimates being $> 300\%$ of the average) (Fig. 1.3). The fraction of the habitat sampled, FRACSAMP, was the most important experimental factor influencing the accuracy of activation energy estimates, as revealed by FRACSAMP consistently being the first split in five of six CART analysis (Fig. 1.3). An exception was when using method M1 (Fig. 1.3 A), the phenomenological likelihood ((equation 1.15, section 1.6.2) for parameter inference, which in general produced relatively inaccurate estimates of activation energy. Therefore, for most methods, sampling $< 0.5\%$ of the habitat leads to the biggest reduction of accuracy (increase in relative error R) in the estimation of activation energy across all experimental factors. Also for the indirect methods which use carrying capacity as a parameter to infer activation energy (methods M2, M4 M6 and M8 in table 1.1) the fraction of habitat searched is the most important experimental factor influencing the accuracy of activation energy estimates (see section 1.6.3 figure 1.7).

After FRACSAMP there was no consistent ordering in the rank importance of the other experimental conditions across the different inference methods (figure 1.3). The number of different temperatures used along a temperature gradient and the number of replicates per experiment were both used for the second split in the classification trees, depending on the inference method used. For the number of replicates, accuracy was significantly lower for experiments with only one replicate than for those with more than one replicate. For example, when the fraction of habitat searched is > 0.005, having at least three replicates instead of only one increases the accuracy of the estimates of activation energy from 16% to 10% error for method M5, from 12% to 6% error for method M7 and from 13% to 6% error for method M9 (figure 1.3 A, B, C and D, respectively). For the number of tempera-

tures, accuracy was significantly lower when just three temperatures were used than when more than three temperatures were used. The number of times in the 15 days period that samples were taken (TIMESAMP) appeared to have the smallest effect, although we expect this was because even the least frequent sampling still included low, medium and high population densities in the time series. Replication also interacts with other factors such as the size of the temperature gradient (TEMPSAMP) to influence the accuracy of the estimates. For example, at low FRACSAMP, increasing the number of temperatures at which experiments are conducted will not increase the accuracy of estimates of activation energy when only one replicate is used per temperature when using indirect methods (figure 1.3 D). However having more temperatures will improve the estimate of activation energy when using a direct method (figure 1.3 F).

Taking into account the observation error in the inference method increased the accuracy of estimates of activation energy when inferring it indirectly for carrying capacity (mean relative error of method M6 of 45% versus mean relative error of M8 is 36%) and growth rate (mean relative error of method M5 is 16% versus mean relative error of M7 is 11%). However, it led to only a minor improvement when inferring activation energy directly (mean relative error of method M9 is 10.6% versus mean relative error of M10 is 10.3%). Estimates of activation energy are generally more accurate when estimated using MCMC parameter inference than using MLE, although sampling a larger fraction of the microcosm can clearly be used to compensate for this (see figure 1.5). Amongst the indirect MCMC methods, more accurate estimates of activation energy were obtained using the inferred growth rate rather than carrying capacity, and accounting for observational error improved these estimates further. These improvements were made with the inevitable cost of computational time (table 1.1).

Figure 1.3 shows the absolute value of the relative error of the estimates of activation energy; however this does not indicate the degree to which the methods are over or underestimating activation energy. This is conveyed in

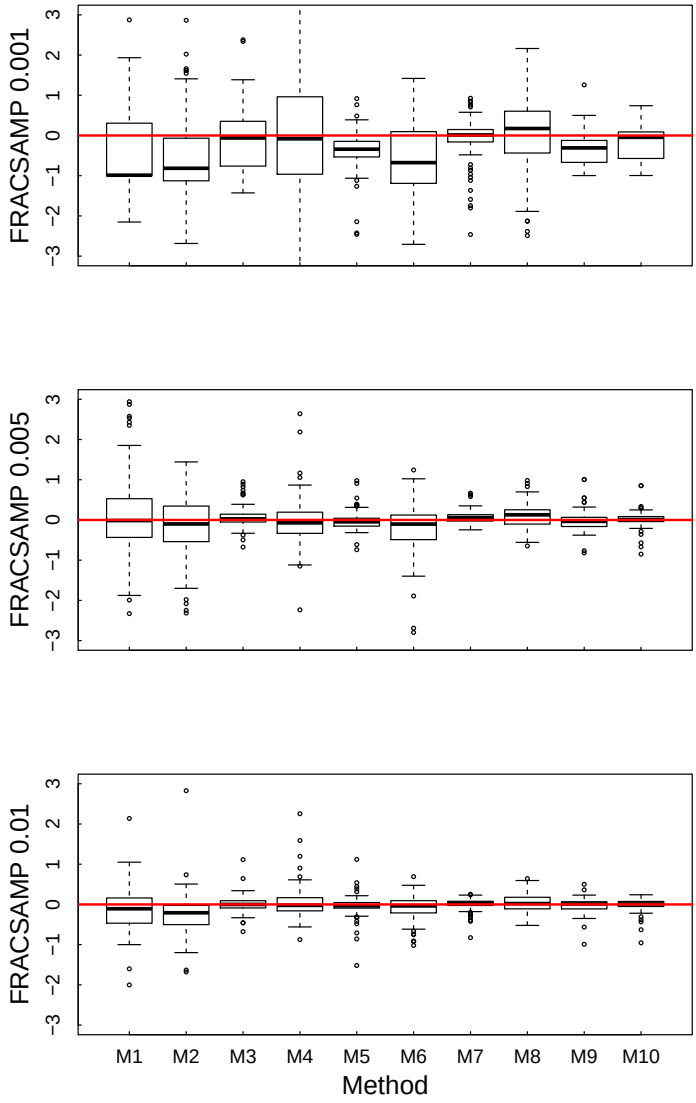

Figure 1.4: The variation in the relative error of each model indicated in table 1 for different FRACSAMP, for experiments with one replicate, for each activation energies used in simulated data ($E_A = 0.2 - 1.2Ev$) for all values of TIMESAMP (5,10,15) and TEMPSAMP (3,6,11). The y axis displays percentage values of relative error. The black lines indicate the medians and the boxes demarcate the 25-75% intervals. The whiskers extend up to one and a half times the inter-quartile range. The red line shows the maximum precision (i.e., estimated value = true value).

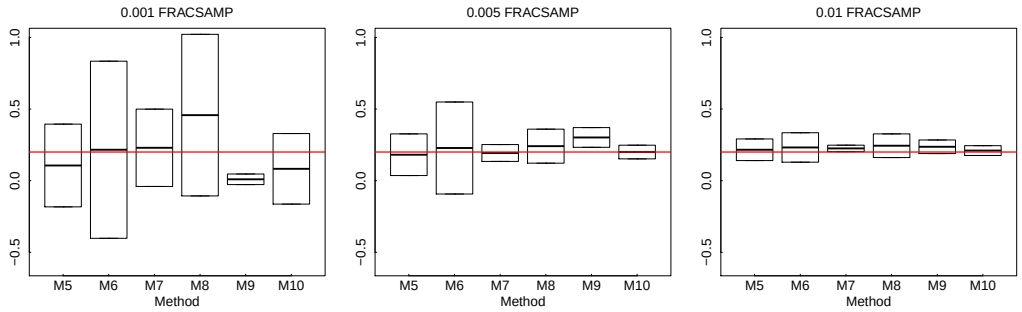

Figure 1.5: Example of different estimates of activation energy (see expression 8) for all the methods based on MCMC (as indicated in table 1) for simulated data (TEMPSAMP = 6, TIMESAMP = 5 REP = 3). The black lines indicate the mean of the estimate and the boxes demarcate the 95% confidence intervals of 1000 samples taken from the Markov chain. The red line shows the real parameter used for simulations ($E_A = 0.2Ev$).

Figures 1.4 and 1.5. These results imply that for most of our methods the true activation energy lies towards the center of the predicted probability distribution for that parameter. An exception is direct inference method M9 in which appears to consistently under predict activation energy at low sampling intensities, which appears to be corrected by taking into account sampling error in method M10. Given the inferior performance of the MLE methods and the dominance of FRACSAMP, we only describe how the precision of estimates is affected by FRACSAMP for the MCMC methods. The most precise estimates of activation energy tend to be obtained using either the direct methods, or the indirect methods on growth rate only with sampling error correction (figure 1.5 M7, M9, M10; the results illustrated in this figure are representative of what we observed for other sets of experimental conditions). In general the most precise estimates were obtained using the direct methods (M9 and M10) which combine information on both growth rates and carrying capacities. Implementing the sampling error correction also tends to increase the precision of the estimated activation energies (figure 1.5). Interestingly direct methods (M9 and M10) are clearly more sensitive to changes in the experimental conditions, as shown by the largest number

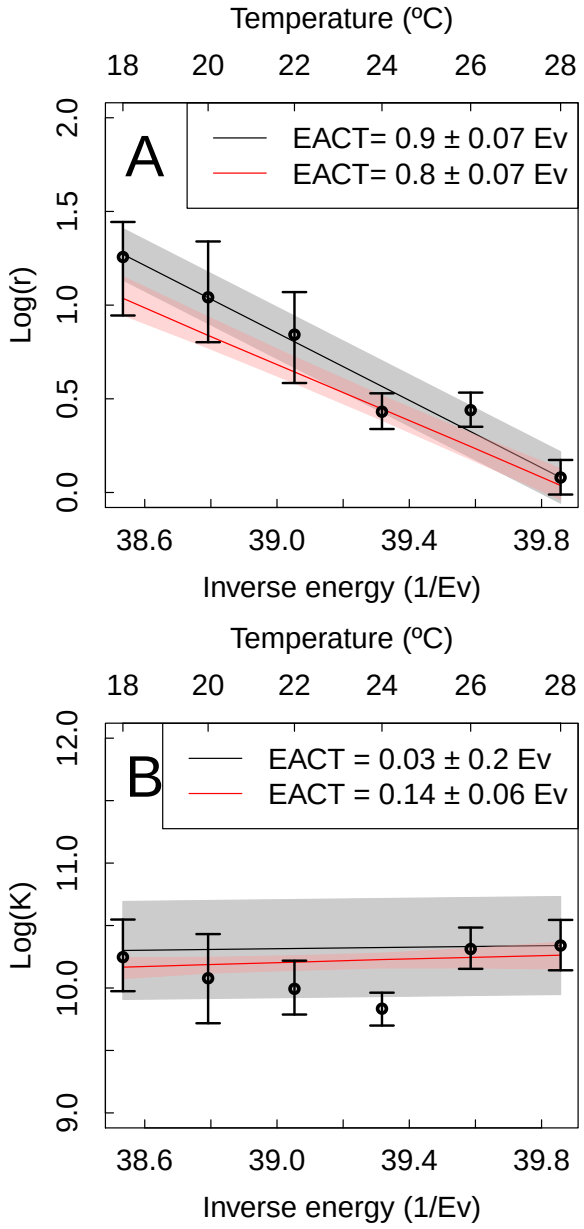

Figure 1.6: Estimates of the logarithm of the growth rate (panel A) and carrying capacity (panel B) of *Paramecium caudatum*. The error bars show the 95% confidence interval of the estimates obtained at each temperature separately. The black continuous line and shaded area represent the estimate of activation energy and the 95% confidence interval of the estimate of activation energy obtained from a weighted linear regression from the values observed at each temperatures (methods M7 for panel A and M8 for panel B, for the methods see table 1.1). The red line and shaded area are the mean and 95% confidence interval of the estimate obtained with method M10 (as in table 1.1) with two different activation energies

of statistically significant branches in the regression trees (figure 1.3 E and F).

When used on real time series data, inferred population growth rate is linearly related to the inverse of temperature, with a negative slope given by the activation energy, as predicted by metabolic theory (Savage *et al.* 2004) (figure 1.6 A). In contrast, the temperature dependence of carrying capacity does not follow the theory (which predicts a positive relationship (Savage *et al.* 2004)), showing no clear directional relationship with temperature (figure 1.6 B). For the best performing methods in our simulation experiments (methods M7, M8 and M10), the direct and indirect methods produce different estimates of activation energy. The estimate for population growth rate from the direct method is slightly lower ($E_A = 0.8Ev$) than the estimate obtained indirectly ($E_A = 0.9Ev$). For the temperature range we considered, this difference leads to the largest contrast between predicted growth rates at $T = 28°C$, where the difference is roughly $1day^{-1}$. Differences in the the mean estimates of activation energy of carrying capacity using direct and indirect methods do not lead to different predicted mean carrying capacities at different temperatures (largely because the estimated activation energy is close to zero). However the precision of those predictions do contrast, for example at $T = 28°C$ the standard deviation of the predicted carrying capacity is approximately 1000 individuals when using the direct method and is approximately 4500 individuals when using the indirect methods. An example of the different estimates obtained with direct and indirect methods at a given temperature ($T = 22°C$) is shown in figure 2.2 D. The activation energy of growth rate measured with the direct method is smaller than the one obtained with indirect methods and has a smaller error.

Applying the phenomenological methods leads to notable differences in the accuracy of the estimates of activation energy for the microcosm experiments. Using indirect method M1 (phenomenological) to estimate activation energy leads to an estimate that is $0.2Ev$ lower than that generated by indirect method M7 ($0.7Ev$ compared to $0.9Ev$, respectively; figure 1.8 A). This difference translates to a difference in predicted growth rate at $T = 28°C$ of

$1.2day^{-1}$. A similar difference is observed when estimating the activation energy of carrying capacity: indirect method M2 (phenomenological) gives an estimate that is $0.2Ev$ higher than that generated by indirect method M8 ($0.03Ev$ compared to $0.2Ev$, respectively; figure 1.8 B). In this example, this could lead to a qualitatively different conclusion about whether carrying capacity is related to temperature, with the phenomenological method implying a positive relationship whereas method M8 implies no relationship.

1.5 Discussion

Our results revealed how experimental factors and parameter inference methods interact to influence the accuracy with which activation energy can be inferred. We found that the fraction of habitat searched is the most important factor in determining the accuracy of the estimates of activation energy. We also provided a list of inference methods from the least to the most accurate, for a set of experimental designs (see figure 1.4), including a classic phenomenological likelihood (Pascual & Kareiva 1996) where no information about demographic stochasticity was included, likelihoods that accounted for demographic stochasticity (Ross, Taimre & Pollett 2006), and likelihoods that accounted for demographic stochasticity and sampling error (Ross, Pagendam & Pollett 2009). Inference methods that included the different sources of stochasticity improved the precision and the accuracy of the estimates of activation energy of at least one order of magnitude, for a given experimental design, especially when the fraction of habitat searched was small. The largest improvement in the accuracy of the estimates was obtained by using a diffusion approximation (Ross, Taimre & Pollett 2006; Ross, Pagendam & Pollett 2009) for continuous time stochastic processes. The use of such approximation enabled us to disentangle different sources of noise (demographic and sampling) and could be extended to more complex models. Another key improvement to the inference was fitting (directly) to all available data simultaneously. Moreover taking into account the sampling error correction in direct methods, where the information of both temperature

dependencies of growth rate and carrying capacity are taken into account, slightly improved the estimate of activation energy. Application of these simulation based findings to real data suggests that although this direct method is more accurate, prior use of the indirect method is useful to reveal the functional form of the temperature dependency.

Comparison of the indirect and direct methods of inference revealed the unique strengths of each approach. Indirect methods are useful to identify the strengths and weaknesses of the different models describing single temperature time series. Once a suitable functional form is implemented, the temperature dependence of ecological parameters can be better inferred using direct methods; yet direct methods could be misleading if applied without having a clear understanding of the outcome of the indirect methods. For example in our study, we based our simulations on a specific exponential function (Arrhenius law) scaled with a single parameter (activation energy). Different functional forms (such as hump shaped functions) would have required a different implementation into direct methods. Similar approaches have been used in other modeling frameworks (Grimm *et al.* 2005; Smith *et al.* 2013) where parameter borrowing between different experiments is used to inform the global parametrization of the model (McInerny & Purves 2011; Sibly *et al.* 2013; Smith *et al.* 2013). The direct approach could be further generalized in more complex models such as food web models (Petchey, Brose & Rall 2010) or stage structured models (Ananthasubramaniam*et al.* 2011). When assessing the performance of different models against data, direct and indirect methods should be combined.

When using direct methods on time series data for *Paramecium caudatum* we found that the estimates of growth rate at each temperature were affected by the estimates of carrying capacity, thus giving "neighbourly advice" (McInerny & Purves 2011) on the temperature dependence of growth rate. The difference in estimation between direct and indirect methods led to large differences in predicted population dynamics (figure 2.2 D). The thermal performance curves of *Paramecium caudatum* have been assessed only us-

ing indirect methods (using growth rate as reference parameter) (Krenek, Berendonk & Petzoldt 2011) and several models have been proposed to capture the temperature dependence of microbial growth (Krenek, Berendonk & Petzoldt 2011; Huang, Hwang & Phillips 2011). We provide a framework to test further the thermal performance of microbial organisms, combining the information of carrying capacity with the information on growth rate. Our methods could be used to compare different thermal performance curves in microbial experiments (Angiletta 2006) and be further tested with different processes such as feeding rates (Rall *et al.* 2009; Englund *et al.* 2011; Fussmann *et al.* 2014) and with different environmental variables such as nutrient concentration (Weisse *et al.* 2002).

The use of stochastic models such as continuous birth and death processes (McKane, & Newman 2004; Black & McKane 2012) provides a probabilistic framework to derive inference schemes from (Ross, Taimre & Pollett 2006; Ross, Pagendam & Pollett 2009) and provides insight into the determinants of population dynamics (Black & McKane 2012). Despite the lack of a mathematical expression for the probability distribution of the populations in our study, the use of approximations, such as the diffusion one, provided an analytical expression for the first two moments of the population probability distribution (Ross, Pagendam & Pollett 2009; Ross 2012). Extending stochastic models to different systems with more than one species is analytically daunting, but numerically feasible. The mechanistic understanding of more complex multispecies models is then limited by their mathematical intractability. When it is not possible to obtain analytical expressions for population probability distributions, the Bayesian framework can be still used with numerical techniques such as particle filters (Ionides 2003; Ionides, Bretó & King 2006) or approximate Bayesian computation (Beaumont 2010). Those methods simulate directly, with a given precision, the likelihood of the model at each iteration of the Markov chain (Hartig*et al.* 2011). Markov Chain Monte Carlo methods are more computationally demanding than classic Maximum Likelihood Estimation, especially when implementing state space models; however they give a more complete estimation of the

probability distribution of the parameters of the model and of their correlation, especially when the distribution of those parameters is not Gaussian.

We chose not to vary N for simplicity in this study although we expect that changes in N to influence our estimates of activation energy in two ways. Firstly, varying N by large amounts (e.g. over an order of magnitude) will significantly change the time the populations take to approach equilibrium, meaning that an adjustment to the sampling design (frequency and intensity) may be needed to obtain a good characterization of the population dynamics. Secondly, the difference between N and K determines the magnitude of demographic fluctuations in the population (see section 1.6.1). As a consequence we expect that differences in N would lead to differences in demographic noise that could influence the precision with which we can estimate activation energy. However the temperature dependence of growth rate and carrying capacity are not dependent on N in our simulation experiments and so we expect that, given an adequate amount of sampling and a sufficiently large temperature range, our conclusions about the effects of likelihood methods and experimental design on estimates of activation energy will be insensitive to our choice of N. Again for simplicity, we assumed that density dependence only influences the probability of births while in reality it commonly influences the probability of both. In section 1.6.1, we give the formulations for the more general birth and death processes in which both birth and death rates depend on N. When combined these lead to more free parameters, but identical formulations for the temperature dependence of population growth rate and carrying capacity, thus our results would be unaffected.

Our methods could improve the development of the ecological theory aimed at understanding the temperature dependence of population rates (Brown *et al.* 2004; Amarasekare & Savage 2012) or inform debates about the precise value of activation energy (Glazier 2006). The use of classic indirect methods can be used as a first step in identifying reasonable functional forms for the temperature dependence of population parameters; as biologists have

extensively done for a variety of taxa (Gillooly *et al.* 2001; Savage *et al.* 2004; Amarasekare & Sifuentes 2012). Different models associated to different functional forms of the rate temperature relations have now been proposed (Brown *et al.* 2004; Knies & Kingsolver 2010; Amarasekare & Savage 2012) and those models, arising from the combination of data and theory, can be further tested using the direct estimation methods we describe here.

One of the remaining conundrums in population and community ecology is about predictive ability. Studies have shown that uncertainty in parameter estimates can preclude predictions of even the direction (increase or decrease) of the effects of a perturbation (Yodzis 1988; Wells, Feldhaar & O'Hara 2014) but also that more accurate estimates will provide better predictions (Novak *et al.* 2011). Our findings support the idea that considerable potential for improved predictive ability lies in improving inference methods, including using quite complex mathematics and fitting algorithms, as well as continuing to use appropriate experimental designs and sampling schemes. The resulting increases in accuracy are likely to be very important, given the documented high sensitivity of model predictions to variation in parameter values.

1.6 Supplementary information

1.6.1 Details on the formulation of the stochastic model

The most general Verhulst-like stochastic birth and death process (Nåsell 2001) is given by Population birth and death rates $B(n)$ and $D(n)$ defined by

$$\begin{aligned} B(n;\theta) &= \theta_1 n\left(1-\theta_2\frac{n}{N}\right), \\ D(n;\theta) &= \theta_3 n\left(1+\theta_4\frac{n}{N}\right). \end{aligned} \tag{1.9}$$

Where n is the population abundance and can take values in the set $0, 1, \cdots, N$. The parameters θ_1 and θ_3 are the per capita intrinsic birth and death rates.

The parameters θ_2 and θ_4 are a measure of the effects of intraspecific competition on birth and death rate respectively and are dimensionless. All the parameters of the model (θ) follow the Arrhenius law (equation 1.2). Process 2.8 is more general than process 1.1 as it includes density dependence also in the death rate. The carrying capacity of process 2.8 is given by $K = N(\theta_1 - \theta_3)/(\theta_1\theta_2 + \theta_3\theta_4)$. N is the population density at which there is zero probability of births (see section **??**) and $N-K$ represent the maximum size of fluctuations associated with demographic stochasticity (Nåsell 2001). In the more general process it is mathematically convenient to scale the strength of density dependence for both birth and death rates with the same parameter, N: the population size at which the probability of births is zero. This makes it more mathematically tractable to calculate the expectation of the population size and the diffusion approximation for population density. In reality, we expect that birth and death probabilities would scale differently with population size, but this could be accommodated by differences in the parameters θ_2 and θ_4 in process 2.8.

The associated stochastic differential equation (SDE) of the process 2.8 is

$$\frac{dn(t)}{dt} = F(n(t);\theta) + \sqrt{H(n(t);\theta)}\frac{dW(t)}{dt}, \tag{1.10}$$

where, from (2.8), we defined two associated functions $F(n;\theta) = B(n;\theta) - D(n;\theta)$ and $H(n;\theta) = B(n;\theta)+D(n;\theta)$, and where W is the standard Wiener process, where $\Delta W(t) = W(t+\Delta t) - W(t)$ has a normal distribution with 0 mean and variance given by Δt (Allen & Allen 2003; Gardinier 2009). The deterministic term in the SDE 2.10 is the classic logistic equation (see equation 3.6 in the main text). It is important to stress that for equations 2.10 and 3.6 $n(t)$ takes continuous values in the interval $[0, N]$. The stochastic term in equation 2.10 is due to random variations in the birth and death rates (demographic stochasticity).

Another way of looking at process 1.1 is to describe it (Gardinier 2009) by

the following master equation

$$\frac{dP(n,t)}{dt} = D(n+1;\theta)P(n+1,t) + B(n-1;\theta)P(n-1,t) - H(n;\theta)P(n,t), \tag{1.11}$$

where $P(n,t)$ is defined as the probability of having n individuals at time t and $H(n;\theta) = B(n;\theta) + D(n;\theta)$. In analogy with chemical kinetics, we call the functions B and D the reaction hazards and H the cumulative hazard of the process (Wilkinson 2006). A detailed mathematical analysis of equation 3.4 is usually intractable, but is straightforward to simulate the time evolution of the system given the rates 1.1. The most common discrete event simulation procedure is known as the *Gillespie algorithm* (Gillespie 1976, 1977).

We obtain process 1.1 from the more general process 2.8 by putting $\theta_4 = 0$ i.e., assuming that intraspecific competition affects only births. Conveniently for our study here, (Ross, Pagendam & Pollett 2009) derived an approximation for the BDP 1.1 which gives the probability of observing a particular number of individuals as a Gaussian distribution with time dependent mean and variance (assuming the maximum population size N is sufficiently large ($N > 1000$; (Ross, Pagendam & Pollett 2009)))

$$P_g(X = x; t, \theta) = \frac{1}{\sqrt{2\pi\sigma(t;\theta)^2}} exp\left(-\frac{x - n(t;\theta)}{2\sigma(t;\theta)^2}\right), \tag{1.12}$$

where the mean of this distribution ($n(t;\theta)$) is given by the solution of the logistic equation 3.6

$$n(t, n_0; \theta(T)) = \frac{K(T)n_0 e^{r(T)t}}{K(T) + n_0(e^{r(T)t} - 1)}, \tag{1.13}$$

and the variance is given by

$$\sigma(t, n_0; \theta(T)) = NM_t^2 \int_0^t H(n(s, n_0, \theta(T))/N; \theta) M_s^{-2} ds, \tag{1.14}$$

where $M_s = \exp\int_0^s B_s ds$ and $B_s = F'(n(s)/N)$, and where $H(n) = B(n) +$

$D(n)$ and $F(n) = B(n) - D(n)$ is the logistic equation 3.6 (Ross, Taimre & Pollett 2006; Ross, Pagendam & Pollett 2009). In section 1.6.2 we describe how we used this approximation to compute the probability of the population being in a particular state given the model parameters.

1.6.2 Likelihoods and inference

Here we describe the likelihood functions used for the inference of activation energy from population time series collected at different temperatures. For each temperature the data is given by $\mathbf{y_{T_k}} = (n_{0,k}, t_0; n_{1,k}, t_1; \cdots; n_{d,k}, t_d)$, where d is the sampling effort (d=TIMESAMP) and $n_{i,k}$ is the number of individuals counted at time t_i and at temperature T_k. A likelihood function whose arguments are the data and the parameters of process 1.1 is associated to each method described in table 1.1.
The first likelihood function we consider is classically used by ecologists in fitting models to time series data (Pascual & Kareiva 1996; Hilborn 1997)

$$L_{phen}(\mathbf{y_{T_k}}; \mathbf{\Theta}(T_k)) = \sum_{i=1}^{d} log\left[P_g\left(X = \Delta n_{i,k}; \mathbf{\Theta}(T_k)\right)\right], \tag{1.15}$$

where $\Delta n_{i,k}(t_i) = log(n_{i,k}) - log(n(t_i - t_{i-1}, n_{i-1,k}; r(T_k), K(T_k))$ is the difference between the logarithms of the observed population densities and the predicted mean densities (see solution 2.13) at every time step, and where the parameters estimated, for every temperature, are $\mathbf{\Theta}(T_k) = (r(T_k), K(T_k), \sigma_k)$ i.e., the growth rate and the carrying capacity together with a variance σ_k. We refer to this as the phenomenological likelihood function because it assigns high likelihood to parameters that capture the phenomenon of logistic population growth without accounting for the effects of the parameters on the demographic stochasticity observed in the population, or for sampling error.

All other likelihood functions incorporate the mathematical derivation of (Ross, Pagendam & Pollett 2009) for the probability of a population having a particular size at a particular time when following the stochastic birth death

process 1.1. If we do not account for sampling error and infer activation energy indirectly then the likelihood function incorporating the probability distribution 2.12 with variance 2.14 is

$$L_1(\mathbf{y_{T_k}}|\theta'(T_k)) = \sum_{i=1}^{d} log\left[P_g(X = n_{i,k}; t_i, \theta(T_k))\right], \tag{1.16}$$

where $\theta'_1(T_k) = log(N/\theta_1(T_k)\theta_2(T_k)) = log(K(T_k)/r(T_k))$, $\theta'_2(T_k) = log(\theta_3(T_k) - \theta_1(T_k)) = log(r(T_k))$, and $\theta'_3(T_k) = log(N/\theta_2(T_k))$. This reparameterization was particularly convenient for the indirect estimation of activation energy because it naturally provides the logarithm of the parameters as in expressions 1.6 and 1.7. More importantly, these specific reparameterisations improved the performance of the inference algorithms, in terms of the rate at which they converged on the correct answer, because they largely removed parameter correlations. We did this for all methods involving likelihoods 4.1, 1.18, 1.19 and 1.20.

One key decision to take while doing inference is whether to account for sampling error. Accounting for sampling error requires us to infer the actual population size at each sampling time given the number of individuals observed in each sample (Cappé, Moulines & Ryden 2005). As we modeled the sampling process using a Poisson distribution, we can include into likelihood function 4.1 a correction account of the form

$$P_f(X = n, \Lambda) = \frac{f\Lambda^{fn}e^{-f\Lambda}}{(fn)!}, \tag{1.17}$$

describing the probability of observing a population of n individuals when the actual size of the population is is Λ and the fraction of habitat searched is f (FRACSAMP = f). We can account for sampling error by adding a correction term of the form 1.17 to likelihood 4.1 which becomes

$$L_2(\mathbf{y_{T_k}}|\overline{\mathbf{y}}_{\mathbf{T_k}}, \theta'(T_k)) = \sum_{i=1}^{d} log\left[P_g(X = \overline{n}_{i,k}; t_i, \theta(T_k))\right] + log\left[P_f(X = n_{i,k}, \overline{n}_{i,k})\right], \tag{1.18}$$

where $\overline{\mathbf{y}}_{\mathbf{T_k}} = (\overline{n}_{0,k}, t_0; \overline{n}_{1,k}, t_1; \cdots; \overline{n}_{d,k}, t_d)$ is a vector of latent variables giving the inferred expected population sizes $\overline{n}_{i,k}$ at time t_i and temperature T_k.

It is straightforward to extend likelihood functions 4.1 and 1.18 to allow activation energy to be inferred directly by incorporating all time series at different temperatures i.e., $\mathbf{Y_T} = \{\mathbf{y_{T_0}}, \mathbf{y_{T_1}}; \cdots; \mathbf{y_{T_q}}\}$. We obtain a more direct method to estimate activation energy by summing likelihoods 4.1 and 1.18 over all possible temperatures of the gradient

$$L_1^D(\mathbf{Y_T}|\theta'_{\mathbf{0}}, \theta_4) = \sum_{k=1}^{q} L_1(\mathbf{y_{T_k}}|\theta'(T_k)), \tag{1.19}$$

$$L_2^D(\mathbf{Y_T}|\overline{\mathbf{Y}}_{\mathbf{T}}, \theta'_{\mathbf{0}}, \theta_4) = \sum_{k=1}^{q} L_2(\mathbf{y_{T_k}}|\overline{\mathbf{y}}_{\mathbf{T_k}}, \theta'(T_k)), \tag{1.20}$$

where q is the size of the temperature gradient ($q = TEMPSAMP$) and $\overline{\mathbf{Y}}_{\mathbf{T}} = \{\overline{\mathbf{y}}_{\mathbf{T_0}}, \overline{\mathbf{y}}_{\mathbf{T_1}}; \cdots; \overline{\mathbf{y}}_{\mathbf{T_q}}\}$ are all the latent variables i.e., the inferred expected population sizes at all temperatures. In this case we infer $\theta'_{\mathbf{0}}$ i.e., the same parameters of the indirect likelihoods at the reference temperature T_0, and a fourth parameter $\theta_4 = log(E_A)$ which provides directly the information about the activation energy of the model. Note that fitting the more general model described by process 2.8 would require different likelihood functions from the one used in models M3-M10. Specifically adding a density dependence on the death rate would produce a different expression for the variance of population density (equation 2.14).

We used two different computational algorithms to estimate the most likely model parameters, one that seeks the maximum likelihood estimate of the parameters (MLE) and one that infers the joint probability distribution of the parameters given the data.

- Details on the MLE optimization methods:

  MLE was performed using function mle2() from the package bbmle in

R (Bolker 2013), using the search method simulated annealing (SANN) (Kirkpatrick, Gelatt & Vecchi 1983) to maximize the likelihood function in parameter space. MLE was performed using two likelihood functions only, 1.15 and 4.1, because attempts to employ MLE to use the other likelihood functions were computationally unfeasible.

Simulated annealing is a stochastic optimization technique which enables to find low cost configurations while still exploring the parameter space (Kirkpatrick, Gelatt & Vecchi 1983). We choose SANN because, among the available optimization methods of the function optim() used by the function mle2(), SANN allowed likelihood estimation to be made where other search algorithms failed to optimize. MLE requires start values from which the search algorithm can begin to search the parameter space. These fitted values where given as normally distributed around the known actual parameters values with variance 1% of the actual value.

- details of the MCMC and filzbach

We inferred the joint probability distributions of the parameters using Markov Chain Monte Carlo sampling with the Metropolis-Hastings algorithm (Chib & Greenberg 1995), implemented using the software Filzbach (Filzbach 2013). We used uniform uninformative priors for all the parameters using likelihoods 4.1, 1.18, 1.19 and 1.20. We did not use this approach with likelihood 1.15 for brevity, after observing the poor parameter estimates using that approach in preliminary analyses. Multiple chains of varying lengths were run initially to check for convergence on a single parameter probability distribution and on the rate of convergence before deciding on a burn in length of 5 million iterations and a sampling period of 5 million iterations. Chains were subsampled every 5000 iterations to remove autocorrelation before analyzing the parameter distributions. Two key features of MCMC sampling make

it attractive for complex, parameter-rich problems. First, as might be expected, the algorithm accepts any change in the parameters that increases the likelihood, but it also probabilistically accepts changes that decrease the likelihood, according to the so-called "Metropolis criterion" (Chib & Greenberg 1995). This latter behavior is particularly important in nonlinear problems, because it allows the algorithm to escape from local maxima of the likelihood, and find the global maximum.

1.6.3 Supplementary results

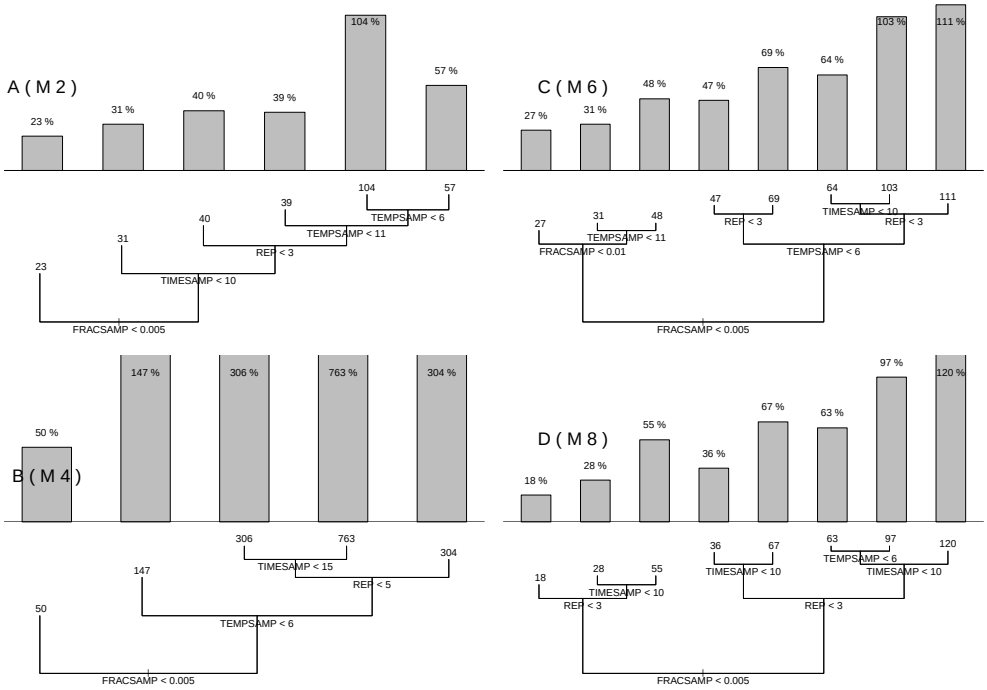

Figure 1.7: The results of the classification and regression tree (CART) analysis (Ripley 2007) of the relative error of the estimates of activation energy. The number at the leaves of the tree indicates the mean percentage value of the relative error of the estimate (see expression 8) over all the simulated experiments, following partitioning of the data in the manor specified by the tree. The threshold above each node indicates the split criterion used to separate the data. To each tree is associate a bar-chart showing the mean percentage value of each leaf. The four panels correspond to four of the models specified in table 1: model M2 (panel A), M4 (panel B), M6 (panel C), M8 (panel D).

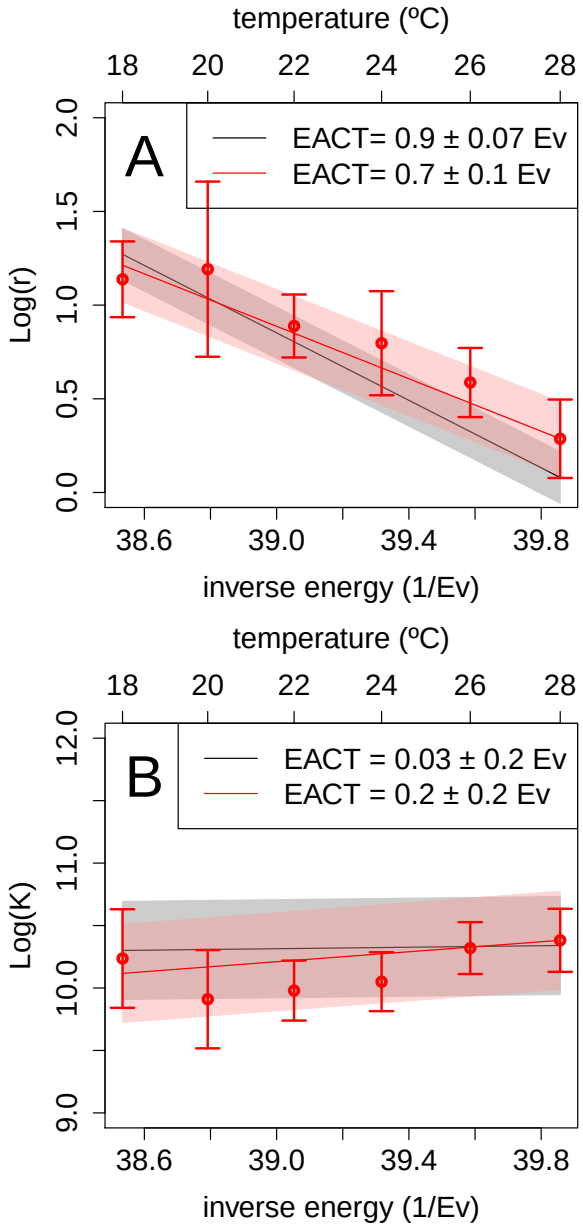

Figure 1.8: Estimates of the logarithm of the growth rate (panel A) and carrying capacity (panel B) of *Paramecium caudatum*. The error bars show the 95% confidence interval of the estimates obtained at each temperature separately using the phenomenological likelihood 1.15. The red continuous line and shaded area represent the estimate of activation energy and the 95% confidence interval of the estimate of activation energy obtained from a weighted linear regression from the values observed at each temperatures (methods M1 for panel A and M2 for panel B, for the methods see table 1). The black line and shaded area represent the estimate of activation energy and the 95% confidence interval of the estimate of activation energy obtained from a weighted linear regression from the values observed at each temperatures obtained using likelihood 18 (methods M7 for panel A and M8 for panel B, for the methods see table 1) as shown in figure 1.6.

Bibliography

Allen, L. J. S. & Allen, E. J. (2003) A comparison of three different stochastic population models with regard to persistence time. *Theoretical Population Biology*, **64**, 439-449.

Allen, A. P., Gillooly, J. F. & Brown, J. H. (2005) Linking the global carbon cycle to individual metabolism. *Functional Ecology*, **19**, 202-213.

Amarasekare, P. & Savage, V. M. (2012) Elucidating the temperature dependence of fitness. *The American Naturalist*, **179**, 178-191.

Amarasekare, P. & Sifuentes, R. (2012) Elucidating the temperature response of survivorship in insects. *Functional Ecology*, **26**, 959-968.

Ananthasubramaniam, B., Nisbet, R. M., Nelson, W. A., McCauley, E. & Gurney, W. S. C. (2011) Stochastic growth reduces population fluctuations in Daphnia-algal systems. *Ecology*, **92**, 362-72.

Anderson, R. M. & May R. M. (1992) *Infectious diseases of humans: dynamics and control*. Oxford University Press.

Angiletta Jr, M. J. (2006) Estimating and comparing thermal performance curves. *Journal of Thermal biology*, **31**, 541-545.

Arrhenius, S. (1889) Über die reaktionsgeschwindigkeit bei der inversion von rohrzucker durch säuren. *Zeitschrift für Physikalische Chemie* **4**, 226-248.

Beaumont, M. A. (2010) Approximate Bayesian Computation in evolution and ecology. *Annual Review of Ecology, Evolution, and Systematics*, **41**, 379-406.

Beveridge, O. S., Petchey, O. L. & Humphries, S. (2010) Direct and indirect effects of temperature on the population dynamics and ecosystem functioning of aquatic microbial ecosystems. *Journal of Animal Ecology*, **79**, 1324-1331.

Black, A. J. & McKane, A. J. (2012) Stochastic formulation of ecological models and their applications. *Trends in Ecology & Evolution*, **27**, 337-345.

Bolker, B. (2013) *Bbmle: Tools for general maximum likelihood estimation.* R package version 1.0-13

Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M. & West, G. B. (2004) Toward a Metabolic Theory of Ecology. *Ecology*, **85**, 1771-1789.

Cappé, O., Moulines, E. & Ryden, T. (2005) *Inference in hidden Markov models.* Springer Series in Statistics XVIII.

Chib, S. & Greenberg, E. (1995) Understanding the Metropolis-Hasting algorithm. *The American Statistician*, **49**, 327-335.

Corkrey, R., Olley, J., Ratkowsky, D., McMeekin, T. & Ross, T. (2012) Universality of thermodynamic constants governing biological growth rates. *PloS One*, **7**, e32003. `http://dx.doi.org/10.1371/journal.pone.0032003`.

De Valpine, P. & Hastings, A. (2002) Fitting population models incorporating process noise and observation error. *Ecological Monographs*, **72**, 57-76.

De'ath, G. & Fabricius, K. E. (2000) Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology*, **81**, 3178-3192.

Dell, A. I., Pawar S. & Savage, V. M. (2010) Systematic variation in the temperature dependence of physiological and ecological traits. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 10591-10596.

Dell, A. I., Pawar, S. & Savage, V. M. (2014) Temperature dependence of trophic interactions are driven by asymmetry of species responses and foraging strategy. *Journal of Animal Ecology*, **83**, 70-84.

Dennis, B., Ponciano, J. M., Lele, S. R., Taper, M. L. & Staples, D. F. (2006) Estimating density dependence, process noise, and observation error. *Ecological Monographs*, **76**, 323-341.

Dennis, B. & Ponciano, J. M. (2014) Density-dependent-state space model for population-abundance with unequal time intervals. *Ecology*, **95**, 2069-2076.

Deutsch, C. A., Tewksbury, J. J., Huey, R. B., Sheldon, K. S., Ghalambor, C. K., Haak, D. C. & Martin, P. R. (2008) Impacts of climate warming on terrestrial ectotherms across latitude. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 6668-6672.

Ebenman, B., Law, R. & Borrvall, C. (2004) Community viability analysis: the response of ecological communities to species loss. *Ecology*, **85**, 2591-2600.

Englund, G., Ohlund, G., Hein, C. L., & Diehl, S. (2011) Temperature dependence of the functional response. *Ecology Letters*, **14**, 914-921.

Ernest, S. K. M., Enquist, B. J., Brown, J. H., Charnov, E. L., Gillooly, J. F., Savage, V. M., White, E. P., Smith, F. A., Hadly, E. A., Haskell, J. P., Lyons, S. K., Maurer, B. A., Niklas, K. J. & Tiffney, B. (2003) Thermodynamic and metabolic effects on the scaling of production and population energy use. *Ecology Letters*, **6**, 990-995.

Filzbach, Microsoft Research Cambridge. `http://research.microsoft.com/en-us/um/cambridge/groups/science/tools/filzbach/filzbach.htm`

Fussmann, K. E., Schwarzmüller, F., Brose, U., Jousset, A. & Rall, B. C. (2014) Ecological stability in response to warming. *Nature Climate Change*, **4**, 206-210.

Gardinier, C. (2009) *Stochastic methods: a handbook for the natural and social sciences*. Springer.

Geritz, S. A. H. & Kisdi, E. (2004) On the mechanistic underpinning of discrete-time population models with complex dynamics. *Journal of Theoretical Biology*, **228**, 261-269.

Gillespie, D. T. (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, **22**, 403-434.

Gillespie, D. T. (1977) Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, **81**, 2340-2361.

Gillooly, J. F., Brown, J. H., West, G. B., Savage, V. M. & Charnov, E. L. (2001) Effects of size and temperature on metabolic rate. *Science* **293**, 2248-2251.

Gillooly, J. F., Charnov, E. L., West, G. B., Savage, V. M. & Brown, J. H. (2002) Effects of size and temperature on developmental time. *Nature*, **417**, 70-73.

Glazier, D. S. 2006. The 3/4-power law is not universal: evolution of isometric, onto-genetic metabolic scaling in pelagic animals. *BioScience*, **56**, 325-332.

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W. M., Railsback, S. F. & DeAngelis, D. L. (2005) Pattern-oriented modeling of agent-based complex systems: lessons from ecology. *Science*, **310**, 987-991.

Hall, A. S. C., Stanford, J. A. & Hauer, F. R. (1992) The distribution of abundance of organisms as a consequence of energy balances along multiple environmental gradients. *Oikos*, **65**, 377-390.

Hartig, F., Calabrese, J. M., Reineking, B., Wiegand, T. & Huth, A. (2011) Statistical inference for stochastic simulation models - theory and application. *Ecology Letters*, **14**, 86-827.

Hilborn, R. (1997) *The ecological detective: confronting models with data*. Princeton University Press.

Huang, L., Hwang, A. & Phillips, J. (2011) Effect of temperature on microbial growth rate-mathematical analysis: the Arrhenius and Eyring-Polany corrections. *Journal of Food Science*, **76**, 553-560.

Huey, R. B. & Kingsolver, J. G. (2011) Variation in universal dependence of biological rates. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 10377-10378.

Ionides, E. L. (2003) Inference and filtering for partially observed diffusion processes via sequential Monte Carlo. *Department of Statistics technical report*, The University of Michigan, Ann Arbor.

Ionides, E. L., Bretó, C., & King, A. A. (2006) Inference for nonlinear dynamical systems. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 18438-18443.

Irlich, U. M., Terblanche, J. S., Blackburn, T. M. & Chown, S. L. (2009) Insect rate-temperature relationships: environmental variation and the metabolic theory of ecology. *The American Naturalist*, **174**, 819-835.

Ives, A. R. (1995) Predicting the response of populations to environmental change. *Ecology*, **76**, 926-941.

Jang, L. & Morin, P. J. (2004) Temperature-dependent interactions explain unexpected responses to environmental warming in communities of competitors. *Journal of Animal Ecology*, **73**, 569-576.

Kirkpatrick, S., Gelatt, C. D. & Vecchi, M. P. (1983) Optimization by simulated annealing. *Science*, **220**, 671-680.

Knape, J. & De Valpine, P. (2012) Fitting complex population models by combining particle filters with Markov chain Monte Carlo. *Ecology*, **93**, 256-263.

Knies, L. & Kingsolver, J. G. (2010) Erroneus Arrhenius: modified Arrhenius model best explains the temperature dependence of ectotherm fitness. *The American Naturalist*, **176**, 227-233.

Krenek, S., Berendonk, T. U. & Petzoldt, T. (2011) Thermal performance curves of Paramecium caudatum: A model selection approach. *European Journal of Protistology*, **47**, 124-137.

Krenek, S., Petzoldt, T. & Berendonk, T. U. (2012) Coping with temperature at the warm edge - patterns of thermal adaptation in the microbial eukaryote Paramecium caudatum. *PloS One*, **7**, e30598. `http://dx.doi.org/10.1371/journal.pone.0030598`

Leary, D. J. & Petchey, O. L. (2009) Testing a biological mechanism of the insurance hypothesis in experimental aquatic communities. *Journal of Animal Ecology*, **78**, 1143-1151.

Leary, D. J., Rip, J. M. K. & Petchey, O. L. (2012)The impact of environmental variability and species composition on the stability of experimental microbial populations and communities. *Oikos*, **121**, 327-336.

Malthus, T. (1798) *An essay on the principle of population.*

McInerny, G. J. & Purves, D. W. (2011) Fine-scale environmental variation in species distribution modelling: regression dilution, latent variables and neighbourly advice. *Methods in Ecology and Evolution*, **2**, 248-257.

McKane, A. J. & Newman, T. J. (2004) Stochastic models in population biology and their deterministic analogs. *Physical Review E*, **70**, 1-19.

McKane, A. J. & Newman, T. J. (2005) Predator-prey cycles from resonant amplification of demographic stochasticity. *Physical Review Letters*, **94**, 1-4.

Monod, J. (1942) *Recherches sur la croissance des cultures bacteriennes.* Hermann, Paris.

Montagnes, D. J. S. & Franklin, D. J. (2001) Effect of temperature on diatom volume, growth rate, and carbon and nitrogen content: reconsidering some paradigms. *Limnology and Oceanography*, **46**, 2008-2018.

Montagnes, D. J. S., Kimmance, S. A., Tsounis, G. & Gumbs, J. C. (2001) Combined effect of temperature and food concentration on the grazing rate of the rotifer Brachionus plicatilis. *Marine Biology*, **139**, 975-979.

Nåsell, I. (2001) Extinction and quasi-stationarity in the Verhulst logistic model. *Journal of Theoretical Biology*, **211**, 11-27.

Novak, M., Wootton, J. T., Doak, D. F., Emmerson, M., Estes, J. A. & Tinker, M. T. (2011) Predicting community responses to perturbations in the face of imperfect knowledge and network complexity. *Ecology*, **92**, 836-846.

O'Connor, M. I., Gilbert, B. & Brown, C. J. (2011) Theoretical predictions for how temperature affects the dynamics of interacting herbivores and plants. *The American Naturalist*, **178**, 626-638.

Pascual, M. A. & Kareiva, P. (1996) Predicting the outcome of competition using experimental data: Maximum Likelihood and Bayesian approaches. *Ecology*, **77**, 337-349.

Petchey, O. L., Brose, U. & Rall, B. C. (2010) Predicting the effects of temperature on food web connectance. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **365**, 2081-2091.

Price, P. B. & Sowers, T. (2004) Temperature dependence of metabolic rates for microbial growth, maintenance, and survival. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 4631-4636.

Rall, B. C., Vucic-Pestic, O., Ehnes, R. B., Emmerson, M. & Brose, U. (2009) Temperature, predator-prey interaction strength and population stability. *Global Change Biology*, **16**, 2145-2157.

Ripley, B. (2007) *Tree: classification and regression trees*. R package version 1.0-26.

Ross, J. V. (2012) On parameter estimation in population models III: time-inhomogeneous processes and observation error. *Theoretical Population Biology*, **82**, 1-17.

Ross, J. V., Pagendam, D. E. & Pollett, P. K. (2009) On parameter estimation in population models II: multi-dimensional processes and transient dynamics. *Theoretical Population Biology*, **75**, 123-132.

Ross, J. V., Taimre, T. & Pollett, P. K. (2006) On parameter estimation in population models. *Theoretical Population Biology*, **70**, 498-510.

Savage, V. M., Gillooly, J. F., Brown, J. H., West, G. B. & Charnov, E. L. (2004) Effects of body size and temperature on population growth. *The American Naturalist*, **3**, 429-441.

Schoolfield, R. M., Sharpe, P. J. H. & Magnuson, C. E. (1981) Nonlinear regression of biological temperature-dependent rate models based on absolute reaction-rate theory. *Journal of Theoretical Biology*, **88**, 719-731.

Sibly, R. M., Grimm, V., Martin, B. T., Johnston, A. S. A., Kulakowska, K., Topping, C. J. & DeAngelis, D. L. (2013) Representing the acquisition and use of energy by individuals in agent-based models of animal populations. *Methods in Ecology and Evolution*, **4**, 151-161.

Smith, M. J., Purves, D. W., Vanderwel, M. C., Lyutsarev, V. & Emmott, S. (2013) The climate dependence of the terrestrial carbon cycle, including parameter and structural uncertainties. *Biogeosciences*, **10**, 583-606.

Thomas, C. D., Cameron, A., Green, R. E., Bakkenes, M., Beaumont, L. J., Collingham, Y. C., Erasmus, B. F. N., De Siqueira, M. F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., Van Jaarsveld, A. S., Midgley, G. F., Miles, L., Ortega-Huerta, M. A., Peterson, A. T., Phillips, O. L. & Williams, S. E. (2004) Extinction risk from climate change. *Nature*, **427**, 145-148.

Turchin, P. (2003) *Complex population dynamics: a theoretical/empirical synthesis*. Princeton University Press.

Van't Hoff, J. H. (1884) *Etudes de dynamique chimique* F. Muller & Co., Amsterdam.

Vasseur, D. A. & McCann, K. S. (2005) A mechanistic approach for modeling temperature-dependent consumer-resource dynamics." *The American Naturalist*, **166**, 184-198.

Vasseur, D. A., DeLong, J. P., Gilbert, B., Greig, H. S., Harley, C. D. G., McKann, K. S., Savage, V., Tunney, T. D. & O'Connor, M. (2014) Increased temperature variation poses a greater risk to species than climate warming. *Proceedings of the Royal Society Series B: Biological Sciences*, **281**, 20132612.

Volkov, I., Banavar, J. R., Hubbell, S. P. & Maritan, A. (2003) Neutral theory and relative species abundance in ecology. *Nature*, **424**, 1035-1037.

Weisse, T. & Montagnes, D. J. S. (1998) Effect of temperature on inter- and intraspecific isolates of Urotricha (Prostomatida, Ciliophora). *Aquatic Microbial Ecology*, **15**, 285-291.

Weisse, T., Stadler, P., Lindström, E. S., Kimmance, S. A. & Montagnes, D. J. S. (2002) Interactive effect of temperature and food concentration on growth rate: A test case using the small freshwater ciliate Urotricha farcta. *Limnology and Oceanography*, **47**, 1447-1455.

Wilkinson, D. (2006) *Stochastic modelling for systems biology*. Chapman & Hall/CRC.

Wells K., Feldhaar, H. & O'Hara, R. B. (2014) Population fluctuations affect inference in ecological networks of multi-species interactions. *Oikos*,

Yodzis, P. & Innes, S. (1992) Body size and consumer-resource dynamics. *The American Naturalist*, **139**, 1151-1175.

Yodzis, P. (1988) The Indeterminacy of Ecological Interactions as Perceived Through Perturbation Experiments. *Ecology*, **69**, 508-515.

Yvon-Durocher, G., Caffrey, J. M., Cescatti, A., Dossena, M., Del Giorgio, P., Gasol, J. M., Montoya, J. M., Pumpanen, J., Staehr, P. A., Trimmer, M., Woodward, G. & Allen, A. P. (2012) Reconciling the temperature dependence of respiration across timescales and ecosystem types. *Nature*, **487** 472-476.

Chapter 2

Predicting the establishment of introduced species: the effect of demographic stochasticity and parameter uncertainty

Gian Marco Palamara, Francesco Carrara, Matthew J. Smith & Owen L. Petchey

2.1 Abstract

Invasive species are becoming a universal component of global change and a serious threat to biodiversity. Therefore, predicting the probability of an introduced species to establish and become invasive is of fundamental importance. Knowledge of the typical strength of competition between resident and introduced species is one of the fundamental ingredients needed to predict the establishment of an introduced species. Many other factors affect the outcome of an introduction, including environmental and demographic stochasticity, Allee effects and spatial effects. Demographic stochasticity becomes particularly relevant when the introduced species is at low density.

We focus on the competitive interaction between residents and potential invader using a stochastic version of the classic competitive Lotka-Volterra (LV) equations. We assess the effect of demographic stochasticity on inferring the interaction parameter and on our ability to predict the probability of an introduced species to establish. Our method is based on a diffusion approximation for the mean and the variance of the population size of the invader and a linearization of the interaction term of the LV model. Using these approximations we provide a modified single species model describing the dynamics of the introduced species and its interaction with the resident species.

We present a novel method to infer competition parameters during the first stages of an introduction, when the invader species remains at low density. We show how having a prior knowledge of the single species demographic parameters can improve the precision of the estimates of the competition parameters of at least one order of magnitude. Finally, we assess how our ability of predicting the establishment of an introduced species depends on demographic stochasticity. Our results provide a first step in disentangling the effects of demographic stochasticity and parameter uncertainty in our ability to predict establishment success.

Key Words: Birth and Death Process, Lotka Volterra Equations, Bayesian Methods, Demographic Stochasticity, Establishment Probability.

2.2 Introduction

Invasive species are one of the major threats to biodiversity (Chapin *et al.* 2000; Cardinale *et al.* 2006; Pereira *et al.* 2010). Due to an increasing invasion rates and pervasive economic costs and environmental damage caused by invasive species, understanding the dynamics of invasions has become a matter of major interest (Vitousek *et al.* 1996; Cacho *et al.* 2006; Simberloff *et al.* 2013). The establishment success of introduced species has been studied intensively for a variety of organisms such as microbes (Litchman 2010), plants (Corbin & D'Antonio 2004; Keane & Crawley 2002; Vilá *et al.* 2011), insects (Kenis *et al.* 2008), fishes (Kolar & Lodge 2002) and birds (Cassey *et al.* 2004; Blackburn, Cassey & Lockwood 2009). It has been argued that establishment success rate of an introduced species strongly depends on both the traits of the introduced species (Williamson & Fitter 1996; Van Kleunen *et al.* 2010) and its interactions with the resident species (Bright 1998; Jeschke & Strayer 2006). Research has shown that the success rate of establishment of an introduced species depends also on several other factors (Williamson & Fitter 1996; Tilman 2004) such as environmental conditions (Beisner *et al.* 2006), the ability of the alien species to adapt to the new habitats (Keane & Crawley 2002; Sax *et. al* 2007; Blackburn, Cassey & Lockwood 2009), the complexity of the food web of the resident community (Romanuk *et al.* 2009; Lurgi *et al.* 2014) and the spatial structure of the landscape (With 2002). Given this variety of interdependent mechanisms, understanding why some introduced species establish while other fail is still a central unanswered question of invasion ecology (Sax *et. al* 2007).

One of the characteristics of an introduction in its earlier stages is the low density of the introduced species. Biological invasions are typically modeled as a traveling wave describing the advancement of the invasion front in space (Giometto *et al.* 2014). The speed of an advancing population wave is deter-

mined by what is happening in the leading edge of the front, where the density of the introduced species is low (Hastings *et al.* 2014). Therefore, several studies have investigated the probability of an introduced species to establish based on density dependent mechanism typically observed at low population density, such as Allee effects (Drake 2004; Taylor & Hastings 2005; Drake & Lodge 2006). Another important mechanism is demographic stochasticity that becomes particularly relevant at low population density, thus at the beginning of an invasion (Drake *et al.* 2006). Demographic stochasticity is caused by random variation among individuals in survival and reproduction. Such variability occurs when all individuals have the same probability of survival and reproduction. Demographic stochasticity differs from environmental stochasticity that refers to the variability in population growth rate caused by temporal differences in the probability of survival and reproduction (Engen, Bakke & Islam 1998). There have been few systematic investigations on the effects of demographic stochasticity in describing the spread of biological invasions (see for example (Snyder 2003; Elliott & Cornell 2013)). Theoretical studies (Romanuk *et al.* 2009; Galiana *et al.* 2014; Lurgi *et al.* 2014) have investigated the effects of the introduction of a species on the stability of a community using deterministic models of population dynamics (Yodzis & Innes 1992; Williams 2008). Such models do not take into account the different stochasticities of population dynamics. Deterministic models are sufficient to describe populations where fluctuations around the average population size are negligible compared to the actual population size (McKane, & Newman 2004), but they ignore the discreteness of populations. The resulting fluctuations at the population level can produce very different dynamics from deterministic models (McKane, & Newman 2005; Ebenman, Law & Borrvall 2004). Demographic stochasticity is an important driver of population dynamics as it is correlated to species interactions (Snyder 2003). Understanding the combined effect of demographic stochasticity and species interaction is, therefore, fundamental to determine the establishment success rate of an introduced species (Lande, Enger & Sæther 2009). Estimating the interaction parameters by adopting solid algorithms, which incorporate demographic stochasticity, is then the key to providing a predictive under-

standing of the establishment of an introduced species in a new community.

An introduced species can interact in several ways with the local pool of resident species (Shea & Chesson 2002). The invader can be a predator, a competitor or a facilitator for one or more of the species in the resident community. These biotic interactions can facilitate or impede the establishment of the introduced species on the resident community according to the strength and the sign of the interaction. Moreover, the interaction term can affect both the birth and the death rate of the introduced species by having the same deterministic (mean field (McKane, & Newman 2004)) effect thus determining the size and the amplitude of the demographic noise (Nisbet & Gurney 1982; Lande, Enger & Sæther 2009). Competitive interaction is particularly relevant in determining the establishment success rate of an introduced species (Case 1990; Tilman 2004; Corbin & D'Antonio 2004). It has been shown, using deterministic models, that the probability of invasion success decreases with the size of the resident community and the average strength of competition between the residents and the introduced species (Case 1990). Often the deterministic assumption of equilibrium has been used as tool to estimate competition parameters (Leslie 1957; Case 1990). Other methods have used transient dynamics to infer competition parameters in both deterministic (Pascual & Kareiva 1996) and stochastic models (Boys, Wilkinson & Kirkwood 2007; Gilioli, Pasquali & Ruggeri 2008). On the other hand, inference methods based on multispecies stochastic models can be computationally expensive (Boys, Wilkinson & Kirkwood 2007; Toni *et al.* 2009) as it is not always possible to have closed analytical forms for the population probability density function and thus for the likelihood function of the model. However, analytical expressions for the mean and variance of population size already exists for single species continuous time stochastic models (Ross, Taimre & Pollett 2006; Ross, Pagendam & Pollett 2009) and are based on diffusion approximations (Pollett 1992). We show that this powerful analytical approach can be applied for two species communities in the context of invasion dynamics. We do this by using another approximation for the introduced species, assuming the resident species re-

main at its carrying capacity during the first stages of the introduction. We propose a method to estimate the interaction parameters, and thus the probability of establishment, based on such approximations. The method applies to an introduced species competing for the same resources of the resident species, and shed light on the relationship between demographic stochasticity, parameters uncertainty and our ability to predict the establishment of the introduced species. The application of this approximations to invasion dynamics is novel in itself.

Using such approximated methods we aim to provide a clear understanding of how parameters uncertainty deriving from demographic stochasticity may affect our ability to predict the outcome of the invasion. More specifically

- We quantify the effect of initial density and demographic traits on our ability to estimate the interaction parameter. In other words, we assess the limits of validity of the approximations used to convert a two species (introduced and resident species) model into a single species model.
- We quantify the effect of prior knowledge of the demographic traits of the introduced species on our ability to infer interaction parameters.
- We quantify the effect of demographic stochasticity on our ability to predict the establishment or not of an introduced species.

2.3 Methods

Continuous time stochastic models of population dynamics such as birth and death processes (BDP) enable to take explicitly into account demographic stochasticity (Nåsell 2001; Black & McKane 2012). The size of the noise caused by demographic fluctuations is given by the difference in birth and deaths rates which in turn are related to the probability of survival and reproduction of the species (Black & McKane 2012). We focus on a continuous time stochastic competitive Lotka Volterra BDP describing the birth and death rates of an introduced species competing for the same resources

of the resident species. In this model, the effect of competition on the demographic noise is expressed by a convenient parameterizations of the birth rate of the introduced species. Using a diffusion approximation (Ross, Pagendam & Pollett 2009), and a linearization of the interaction term we can obtain a closed form for the mean and variance of the population size of the introduced species. We propose then an inference framework based on such approximations showing how the size of the demographic processes and the corresponding parameter's uncertainty affects our ability to predict establishment success of the introduced species.

2.3.1 The Model

The classic competitive Lotka Volterra model (Lotka 1920) is a set of differential equations given by

$$\frac{dn_I}{dt} = r_I n_I \left[1 - \left(\frac{n_I + \alpha_{IR} n_R}{K_I}\right)\right], \tag{2.1}$$

Where $n_I(t)$ and $n_R(t)$ are the population densities of the introduced and the resident species at time t, r_I and K_I are the growth rate and carrying capacity of the introduced species and α_{IR} is the per capita effect of the resident on the introduced species. A symmetric equation describes the population dynamics of the resident species (Lotka 1920). Equation 2.1 can be generalized by a stochastic BDP represented by the following birth and death rates for the introduced species

$$\begin{aligned} B(n_I) &= \lambda_I n_I \left[1 - \left(\frac{n_I + \alpha_{IR} n_R}{N_I}\right)\right], \\ D(n_I) &= \mu_I n_I, \end{aligned} \tag{2.2}$$

where λ_I and μ_I are the per capita intrinsic birth and death rates of the invader ($r_I = \lambda_I - \mu_I$) and N_I is the population size at which there is zero probability of births for the introduced species. In process 2.2 the carrying capacity of the introduced species K_I is explicitly separated from its

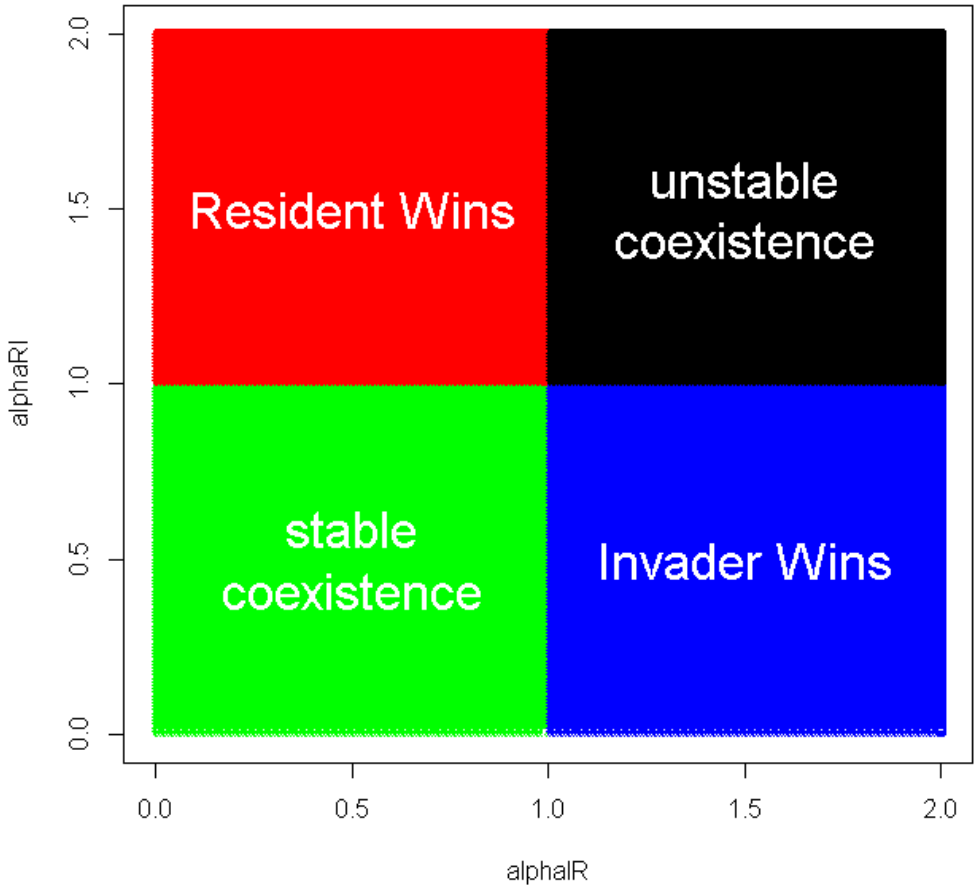

Figure 2.1: Possible determinisitic outcomes of equations 2.1 when $r_I = r_R = 1day^{-1}$ and $K_I = K_R = 10000$ individuals.

maximum population size N_I which in turn is a physical limit related to the size of the habitat. N_I and K_I are related by the intrinsic birth and death rates which define the maximum size of the demographic fluctuations at equilibrium ($K_I = r_I N_I/\lambda_I$). Demographic stochasticity becomes relevant for population dynamics when population size is small compared to the maximum population size ($n_I < \sqrt{N_I}$) as is the case during the first stages of an invasion. For simplicity, we assume inter- and intra-specific competition affect only the birth rate of the introduced species. Process 2.2 is more general than equations 2.1 as it describes the time evolution of the probability distribution of having n_I individuals of the introduced species at time t. Only the mean of such probability distribution is given by the solution of equations 2.1 (for details of the general Lotka Volterra BDP see supplementary information in section 2.6).

Assuming the resident species remains at its single species carrying capacity (K_R) we can reduce the process 2.2 to a single species BDP for the introduced species, with a modified birth rate given by

$$\begin{aligned} B(n_I) &\simeq \overline{\lambda}_I n_I \left(1 - \frac{n_I}{N_I}\right), \\ D(n_I) &= \mu_I n_I, \end{aligned} \tag{2.3}$$

where

$$\overline{\lambda_I} = \lambda_I \left(1 - \frac{\alpha_{IR} K_R}{N_I}\right) \tag{2.4}$$

is the modified intrinsic birth rate of the introduced species which takes into account the effect of the interaction between the resident and the introduced species. The approximated process 2.3 is a convenient version of the stochastic logistic equation (Nåsell 2001; Ross, Taimre & Pollett 2006) with a modified growth rate ($\overline{r}_I = \overline{\lambda}_I - \mu_I$) and carrying capacity ($\overline{K}_I = \overline{r}_I N_I/\overline{\lambda}_I$). It can be shown that as long as the resident species remains close to its original carrying capacity, the introduced species has an intrinsic birth rate given by expression 2.4 (see supplementary information in section 2.6).

To conduct parameter inference we need a mathematical function defining the probability of a set of parameters given the data i.e., the likelihood function. The likelihood function takes as arguments the data and $\theta = (log(r_I), log(K_I), log(N_I), log(K_R), \alpha_{IR})$ i.e., the parameters of process 2.3. Time series data for the population of the introduced species (hereafter labeled as $n_I \equiv n$) is given by $\mathbf{y} = (n_0, t_0; n_1, t_1; \cdots; n_d, t_d)$, where d is the sampling effort and n_i is the number of individuals counted at time t_i. Conveniently for our study, (Ross, Pagendam & Pollett 2009) derived a diffusion approximation for the BDP 2.3 which gives the probability of observing a particular number of individuals as a Gaussian distribution with time dependent mean and variance (assuming the maximum population size N_i is sufficiently large (Ross, Pagendam & Pollett 2009))[1]. The likelihood function is thus given by

$$L(\mathbf{y}|\overline{\mathbf{y}}, \theta) = \sum_{i=1}^{d} log\left[P_g(X = \overline{n}_i; t_i, \theta)\right] + log\left[P_f(X = n_i, \overline{n}_i)\right], \tag{2.5}$$

where P_g is a normal distribution with mean and variance parameterized with θ and P_f is a Poisson distribution describing the probability of observing a population of $\overline{n}$ individuals when the actual size of the population is is n and the fraction of habitat searched is f ($f < 1$). The first term of the likelihood 2.5 describes the demographic process 2.3 and the second term describes the sampling error associated to any experiment or field survey.

2.3.2 Simulations

We fixed both the growth rate and the carrying capacity of the resident and the introduced species at at[2] $r_I = r_R = 1 days^{-1}$ and $K_I = K_R = 10000$ individuals. In this way we can obtain all the possible deterministic behaviors

[1]The same approximation can be applied to the BDP 2.2 (Ross, Pagendam & Pollett 2009) but it is not possible to obtain a closed form for the mean and the variance of population size.

[2]the unit of measure of growth rate can be adjusted to the generation time of the invader species.

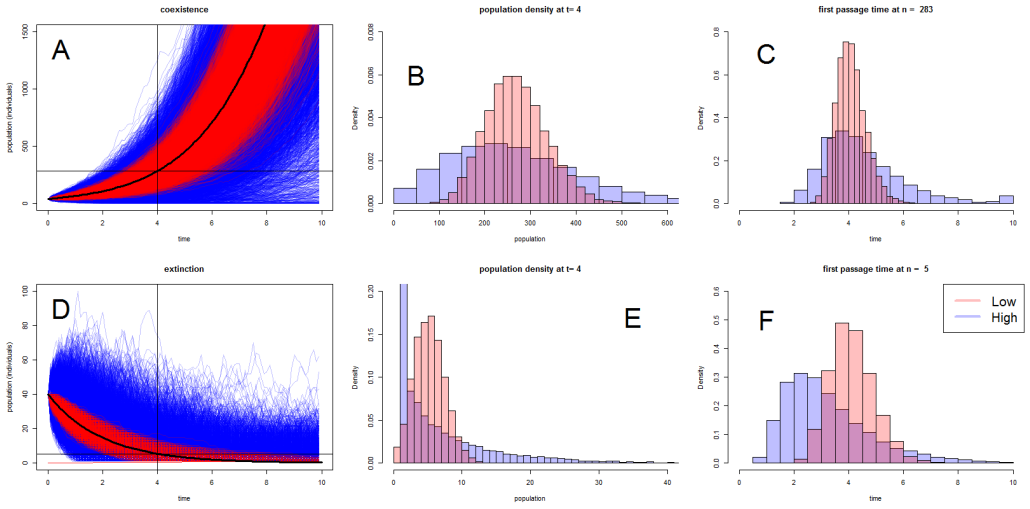

Figure 2.2: 10000 replicates of population dynamics described by process 2.2, where the resident species starts at its single species carrying capacity ($K_R = 10000$ individuals) and the introduced species starts at initial condition $n_I = 40$ individuals. We show two possible dynamics, when the introduced species establishes (coexistence) ($\alpha_{IR} = 0.5$) (Panel A) and when the introduced species goes extinct ($\alpha_{IR} = 1.5$) (Panel D). The two colors refer to low demographic noise (red, $\delta = 1$) and high demographic noise (blue, $\delta = 4$). The black curved lines show the deterministic dynamics predicted by equation 2.1. In panels B and E we show the population distribution of the introduced species at $t = 4days$ (i.e. when the resident abundance is deterministically above 99% of its single species carrying capacity). In panel C and F we show the distribution of the times at which $n_I(t = 4days)$ is reached.

of the Lotka Volterra equations by varying the interaction parameters in the range $[\alpha_{IR}, \alpha_{RI}] \in [0, 2] \times [0, 2]$ (see fig 2.1). Any other choice of parameter values can be conducted to those values by an appropriate renormalization of the competition coefficients and the carrying capacities (see supplementary information in section 2.6). We then fix $\alpha_{RI} = 0.5$ so that all possible dynamics are reduced to establishment of the introduced species i.e., coexistence ($\alpha_{IR} \in [0, 1]$) or not establishment of the introduced species i.e., extinction ($\alpha_{IR} \in [1, 2]$). For $\alpha_{IR} = 1$ there is a bifurcation in the deterministic model, from establishment (coexistence) to not establishment (extinction) of the introduced species. Given the above constraints on the parameters of the model we can now simulate the same deterministic dynamics with different demographic noise by varying a single parameter δ where $\lambda_I = \lambda_R = \delta + 0.5$, $\mu_I = \mu_R = \delta - 0.5$ and $N_I = N_R = 10000(\delta + 0.5)$ and keeping growth rate and carrying capacity fixed. For simplicity, we changed in the same way the demographic parameters of the resident and the introduced species.

We simulated the process 2.2 using the Gillespie algorithm (Gillespie 1976, 1977) producing continuous time series recording the exact times of individual birth and death events (see fig 2.2 for examples) along a gradient of demographic stochasticity given by $\delta \in [1, 10]$. As initial conditions we always set the resident at its single species carrying capacity ($K_R = 10000$) and the introduced species at $n_0 < 100$ thus ensuring that for any interaction coefficient, and for every $t < 5days$, the resident remained (deterministically) at at least 99% of K_R. The continuous time series described in figure 2.2 were then sampled at discrete times (twice a day for the first five days i.e., $d = 10$ in likelihood 2.5), reproducing the time series data obtained in real experiments. We set the sampling error to $f = 0.25$ reproducing a search effort of the 25% of the habitat. We tested the field situation (one replicate for each parameter combinations) and the laboratory situation (5 replicates for every parameter combinations). We do not explore different experimental designs as in (Palamara *et al.* 2014) a systematic investigation of the effects of the experimental design on the estimation if single species parameters has already been performed. We inferred the joint probability distributions of

the parameters θ from the likelihood 2.5 using Markov Chain Monte Carlo sampling with the Metropolis-Hastings algorithm (Chib & Greenberg 1995), implemented using the software Filzbach (Filzbach 2013). We measured the precision and the relative error of the estimate of the interaction parameter, given by the standard error of the estimate and the absolute distance between the estimated value and the "real" value used for simulations. Accuracy of the estimate can be seen as the inverse of the relative error. Precision and accuracy in turn determines our ability to predict the outcome of the interaction i.e. establishment or not establishment of the introduced species.

We tested our ability to estimate α_{IR} under different scenarios representing our knowledge of single species parameters. First we fixed all the single species parameters to their real value in likelihood 2.5, reproducing an unrealistic situation of perfect knowledge of the invader properties in the resident's habitat. We use this scenario to test the effect of the initial condition (i.e. the initial density of the introduced species) and of demographic stochasticity on the precision and accuracy of the estimate of α_{IR} for both a field (one replicate) and a laboratory (five replicates) situation (see figure 2.3). We then relaxed the assumptions on the knowledge of the single species parameters by putting priors on them (figure 2.4). We put priors on the carrying capacity of the introduced species only, reproducing the scenario in which we know very well the demographic parameters of the introduced species but we have less knowledge on how the introduced species would grow in the same environmental conditions of the resident species. We then put priors on both K_I and N_I reproducing a perfect knowledge of the growth rate of the introduced species only, and a reduced knowledge of both its carrying capacity and demographic parameters (figure 2.4). Finally we put priors on all the single species parameters. We always fix the resident carrying capacity to its real value and use a flat uninformative prior for α_{IR}. Lastly we computed the probability of predicting the establishment of the introduced species according to the interaction parameter. By varying α_{IR} and δ and measuring how the probability of having a wrong prediction i.e. of predicting an establishment when the introduced species doesn't establish and vice

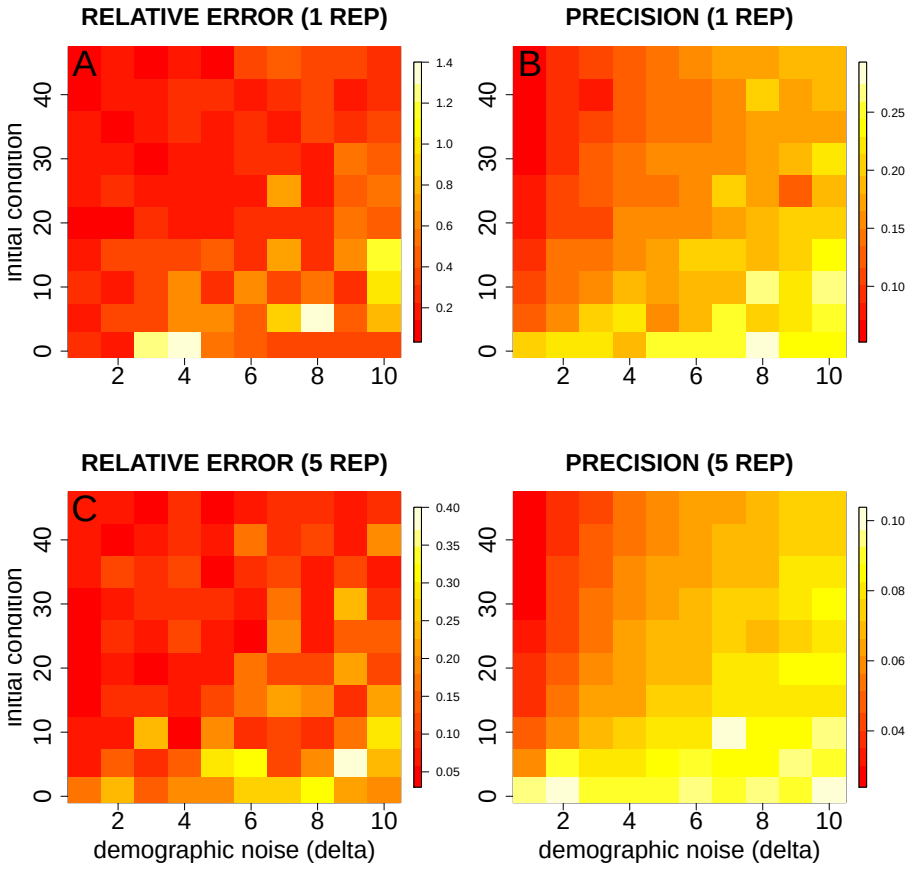

Figure 2.3: Relative error (Panel A and C) and precision (Panel B and D) of the estimate of the interaction parameter ($\alpha_{IR} = 0.5$) as a function of the initial condition of the introduced species and of demographic noise (δ). The estimates are obtained with simulated data generated by the process 2.2 using likelihood 2.5. The data where simulated with 1 replicate (panels A ad B) or 5 replicates (panels C and D), 10 observations ($d = 10$) and a fraction of habitat searched of $f = 0.25$ for all the points in the plots. All single species parameters were fixed to their real value. Each of the points on the plot shows averages over ten realizations of the same numerical experiment.

versa (figure 2.5). We computed this probability only for the field situation i.e. for one replicate.

2.4 Results

The dynamics arising from the same deterministic invasion but with different demographic parameters can be quite distinct (see figure 2.2). A higher demographic noise (higher δ) flattens both the distribution of population densities (see figure 2.2 C) and the first passage time's distribution (see figure 2.2 D) i.e. the distribution of times needed to reach a fixed population threshold. An introduced species with higher birth and death rates (higher δ) will experience more fluctuations around its expected mean trajectory and thus will be more likely to go extinct during the first stages of the introduction thus reducing the probability of establishment. This is not surprising and has already been shown by more phenomenological models (Snyder 2003). Given the flattened population distributions of figure 2.2 it becomes also harder to estimate correctly the interaction parameters when demographic noise increases. The precision and the relative error of the estimates of the interaction parameter are both reduced with increasing demographic noise (see figure 2.3). At the same time the initial density of the introduced species can balance the effect of the demographic noise: when the initial density of the introduced species is higher it becomes easier to estimate correctly the interaction parameter and thus the establishment probability, despite the demographic noise (see figure 2.3). There is , however, more variability in the relative error of the estimates of α_{IR} at higher demographic noise and at lower initial density of the introduced species (see figure 2.3 A and C). In the one replicate situation the relative error of the estimate of α_{IR} can be up to 150 % (figure 2.3 A) together with a standard error of up to 30% (figure 2.3 B). In an experimental situation (5 replicates of the same invasion) relative error of the estimate of α_{IR} (figure 2.3 C) and its standard error (figure 2.3 D) are reduced by one third.

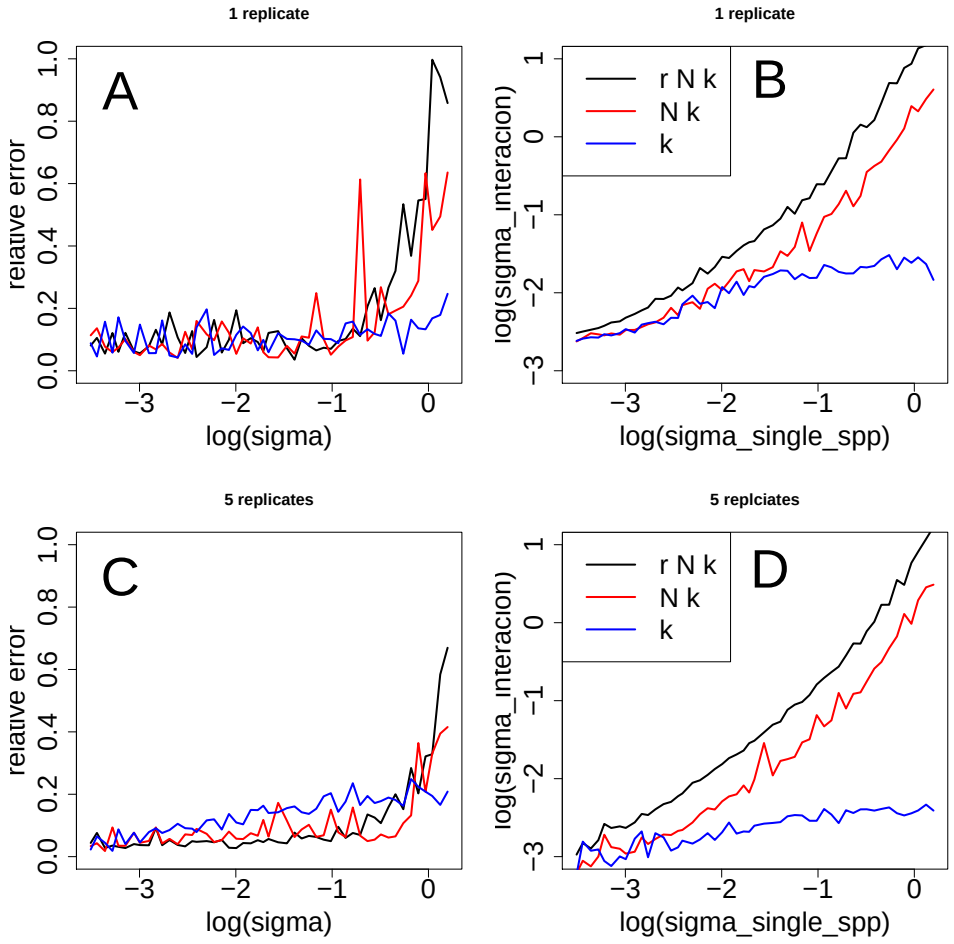

Figure 2.4: Accuracy and precision of the estimate of the interaction parameter ($\alpha_{IR} = 0.5$) as a function of the variance of the Gaussian priors used for the single species parameters. The different colors refer to cases where one (blue) two (red) or three (black) single species parameters are set as priors while the others are fixed to the real value used to simulate the data. The estimates are obtained with simulated data generated by the process 2.2 using likelihood 2.5. Each of the points on the plot shows averages over 20 realizations of the same numerical experiment.

Another important source of error for the estimation of the probability of establishment success is the prior knowledge of the single species parameters; from figure 2.4 it is evident that only being uncertain on the carrying capacity of the introduced species on the resident's habitat and knowing the demographic parameters of the introduced species will keep the accuracy of the estimate of the interaction parameter below 20 % and its relative error relatively constant at a low value (13%) (see figure 2.4). On the other hand, when there is uncertainty also on the demographic parameters of the introduced species, the precision of the estimate of α_{IR} is correlated to the precision of our knowledge of the demographic parameters, resulting in an increase in the variance of the estimate of the interaction parameters as the variance of the single species parameters priors increases (see figure 2.4 B and D). The slope of this variation is slightly reduced when we know the growth rate of the invader but we have priors on both N and K thus on the knowledge of birth and death rates of the invader separately. There is more variability in the relative error of the estimate of α_{IR} in the one replicate situation (figure 2.4 A) compared to the laboratory situation (figure 2.4 B). In the one replicate situation there is, however, a slight increase in the standard error of the estimates of the interaction parameter (figure 2.4 B) compared to the laboratory situation (figure 2.4 B). Figure 2.4 shows that an accurate knowledge of the demographic parameters of the introduced species becomes very important for having a precise estimate of the interaction parameter. Having priors on the single species parameters of the introduced species affects mostly the standard error of the estimates of the interaction parameters and have little effect on the relative error of the estimate.

Finally, we show how demographic noise can reduce the probability of predicting the establishment of an introduced species. Figure 2.5 shows an asymmetry of the effect of demographic noise on the probability of predicting the deterministic establishment of an introduced species and the probability of predicting the failure to establish. In other words, if the introduced species establishes deterministically ($\alpha_{IR} < 1$) the demographic noise can reduce our ability to predict the establishment, especially when the interaction

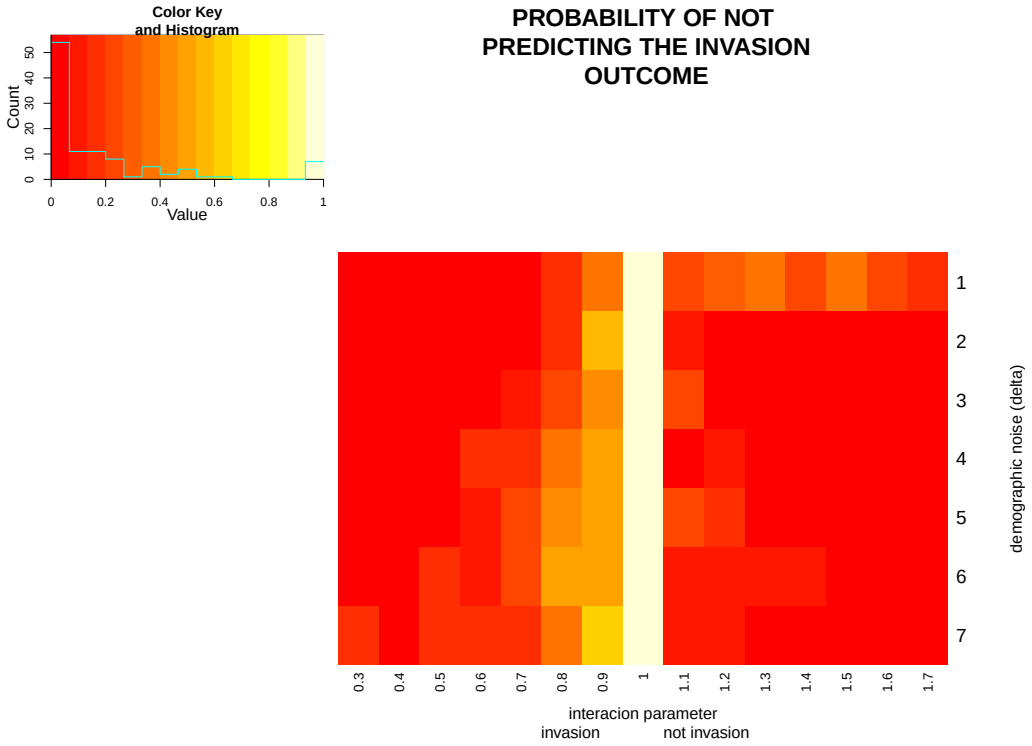

Figure 2.5: Probability (figure 2.5) of not predicting the invasion outcome as a function of the demographic noise (δ) and the value of the interaction parameter (α_{IR}). The estimates are obtained with simulated data generated by the process 2.2 with initial condition $n_0 = 40$, using likelihood 2.5, when all single species parameters are fixed to the real values used to simulate the data. Each of the points on the plot shows averages over 70 realizations of the same numerical experiment.

parameter is close to the deterministic bifurcation ($\alpha_{IR} = 1$). This means that the deterministic model predicts establishment but that demographic stochasticity prevents it. Instead, if the introduced species is not establishing ($\alpha_{IR} > 1$), the demographic noise doesn't effect our ability to predict the failure to establish. However at low demographic noise there is still a finite probability of predicting establishment when the introduced species doesn't establish (see figure 2.2). This is because the introduced species experiences fewer fluctuations around its deterministic trajectory.

2.5 Discussion

We presented a novel method to infer competition parameters from the first stages of an invasion, taking into account the effect of demographic stochasticity. The results from this study provide a clear understanding of how demographic stochasticity biases our predictive understanding of invasion dynamics. In fact, demographic stochasticity not only affects the probability of establishment (Snyder 2003; Elliott & Cornell 2013), but also our ability to infer such probability. Our method is, therefore, specifically designed to take into account the demographic noise when inferring interaction parameters and their effects on invasion dynamics. We shown how our inference method, based on continuous time stochastic models, can become a powerful tool to infer the probability of establishment of an introduced species. Our method represents an advance in the recently developed invasion science (Simberloff 2011; Simberloff *et al.* 2013)

The method we presented can be tested on both time series data from real systems such as lakes (Drake *et al.* 2006) and experimental systems such as microcosms experiments (Drake *et al.* 2011). In fact, our method is suited for predicting invasion success when the single species demographic parameters of the invader are known from previous experiments. Microcosms have played a fundamental role in shaping ecological theory (Drake, Huxel & Hewitt 1996; Drake *et al.* 2011) and could be used to infer the demographic parameters of introduced species. We could test our method in several realistic situations for which we have or not have prior information on the demographic traits of the introduced species as we already did for simulated data (see figure 2.4). Having information on the different stages of the invasion of an introduced species would enable to test the predictions of establishment success that our method provides.

Another relevant extension of our study would be a systematic investigation of the experimental design used for inferring the interaction parameters. A similar investigation has already been performed for single species time series

data (Palamara *et al.* 2014) and the fraction of habitat searched has been found to be the most important factor determining the precision and the accuracy of single species parameters. We, therefore, believe that a similar result will be found for the double species case. Such an investigation would enable to design experiments to infer competition parameters between different species. The only requirement needed to infer the interaction parameters with our method is that one of the two competing species starts at its single species carrying capacity. Then, in order to infer all pairwise interactions, resident and invader can be switched into two separate experiments. Being based on transient dynamics our method can be also used to design experiments to detect density dependent effects and environmental effects on competition and the results compared with already existing estimation methods (Leslie 1957; Pascual & Kareiva 1996; Boys, Wilkinson & Kirkwood 2007; Toni *et al.* 2009).

The mathematical approximations developed in this study also provide a possible framework where invasion theory can be further developed. Process 2.3 and the correspondent likelihood function 2.5 can be improved in order to include environmental and spatial effects (Drake *et al.* 2006) and other mechanisms such as the Allee effect (Drake 2004; Taylor & Hastings 2005). As refined approximation techniques have been developed to predict the extinction risk (Ovaskainen & Meerson 2010), based on the low density of endangered species, similar methods can be developed to predict the invasion risk of alien species. Furthermore, the approximations developed for the competitive Lotka-Volterra model could also be extended to other types of interactions such as predator prey interaction (Palamara *et al.* 2013). The effect of introduced species usually with theoretical models of food webs with deterministic dynamics (Lurgi *et al.* 2014) can then be completed adopting our perspective of approximating the effect of interactions locally around one or more focal species.

Invasion science has recently undergone a shift from primary focus on impacts on particular species to cumulative impacts on ecosystems (Simberloff *et al.*

2013; Lurgi *et al.* 2014). We propose to bridge the gap between species-centric and community perspective, via approximating the effect of interspecific interactions between resident and introduced species. The use of continuous time stochastic models has been proposed for many biological systems, such as evolutionary processes (Kimura 1964) and epidemics (Alonso, McKane & Pascual 2007; Black, & McKane 2010), but their application in ecology is limited (Ross, Taimre & Pollett 2006). In this paper we presented simple example of how such powerful mathematical tools can be used to improve our predictive understanding of invasion dynamics, and propose a general framework that can be extended to more complex ecological systems and dynamics.

2.6 Supplementary information

In this section we describe a general birth and death processes describing the dynamics of a community of S species competing for the same resources. The process is given by two vectors of population birth and death rates $\mathbf{B}(\mathbf{n})$ and $\mathbf{D}(\mathbf{n})$ defined by

$$\begin{aligned} B_i(\mathbf{n}) &= \lambda_i n_i \left(1 - \frac{\sum_j^S \alpha_{ij} n_j}{N_i}\right), \\ D_i(\mathbf{n}) &= \mu_i n_i \left(1 + \frac{\sum_j^S \beta_{ij} n_j}{N_i}\right), \end{aligned} \tag{2.6}$$

where $0 \leq n_i \leq N_i$ is the (integer) number of individuals of species i ($1 \leq i \leq S$), N_i is population size at which species i has zero probability of births, λ_i and μ_i are the birth and death probabilities in the absence of density dependence, respectively (units: day^{-1}). Matrices α and β contain the per capita effects of competition on births and deaths respectively. The diagonal part of α and β contains the intra-specific competition coefficients, while the off-diagonal parts contain the interspecific competition coefficients. Process

2.6 is a stochastic version of the classic competitive Lotka-Volterra equations

$$\frac{dn_i}{dt} = r_i n_i \left(1 - \frac{\sum_j^S A_{ij} n_j}{K_i}\right), \tag{2.7}$$

where $r_i = \lambda_i - \mu_i$ is the per capita growth rate of species i, $K_i = r_i N_i/\lambda_i$ is the carrying capacity of species i and $A_{ij} = \alpha_{ij} + \mu_i \beta_{ij}/\lambda_i$ is the per capita effect of species j on species i. The deterministic model 2.7 can be seen as a first approximation of the stochastic model 2.6. Setting $\beta = 0$ (i.e. there is no effect of competition on the death rate of any species) and setting the number of species to two ($i \equiv I$ and $j \equiv R$) we obtain the simpler process 2.2 used to describe the dynamics of an invasion.

Assuming $n_j \equiv K_j$ (i.e. species j remains at fixed density K_j) process 2.6 can be linearized into a general Verhulst-like stochastic birth and death process (Nåsell 2001) for species i given by population birth and death rates $B_i(n_i)$ and $D_i(n_i)$ defined by

$$\begin{aligned} B_i(n_i) &= \overline{\lambda}_i n_i \left(1 - \frac{n_i}{N_i}\right), \\ D_i(n_i) &= \overline{\mu}_i n_i \left(1 + \frac{n_i}{N_i}\right). \end{aligned} \tag{2.8}$$

The modified intrinsic birth and death rates are given by

$$\begin{aligned} \overline{\lambda}_i &= \lambda_i \left(1 - \frac{\alpha_{ij} K_j}{N_i}\right), \\ \overline{\mu}_i &= \mu_i \left(1 + \frac{\beta_{ij} K_j}{N_i}\right), \end{aligned} \tag{2.9}$$

where the only assumption done is that the second order terms ($O(\frac{1}{N_i^2})$) are negligible. Note that when $\beta = 0$ we obtain expression 2.4.

The associated stochastic differential equation (SDE) of the process 2.8 is

$$\frac{dn_i(t)}{dt} = F_i(n_i(t)) + \sqrt{H_i(n_i(t))}\frac{dW(t)}{dt}, \tag{2.10}$$

where, from 2.8, we defined two associated functions $F_i(n_i) = B_i(n_i) - D_i(n_i)$ and $H_i(n_i) = B_i(n_i) + D_i(n_i)$, and where W is the standard Wiener process, where $\Delta W(t) = W(t + \Delta t) - W(t)$ has a normal distribution with 0 mean and variance given by Δt (Allen & Allen 2003; Gardinier 2009). The deterministic term in the SDE 2.10 is the classic logistic equation (Nåsell 2001). The stochastic term in equation 2.10 is due to demographic stochasticity.

Another way of looking at process 2.8 is to describe it (Gardinier 2009) by the following master equation

$$\frac{dP(n_i,t)}{dt} = D_i(n_i+1)P(n_i+1,t) + B_i(n_i-1)P(n_i-1,t) - H_i(n_i)P(n_i,t), \tag{2.11}$$

where $P(n_i, t)$ is defined as the probability of having n_i individuals of species i at time t. In analogy with chemical kinetics, we call the functions B_i and D_i the reaction hazards and H_i the cumulative hazard of the process (Wilkinson 2006). A detailed mathematical analysis of equation 3.4 is usually intractable, but is straightforward to simulate the time evolution of the system given the rates 1.1. The most common discrete event simulation procedure is known as the *Gillespie algorithm* (Gillespie 1976, 1977).

Conveniently for our study here, (Ross, Pagendam & Pollett 2009) derived an approximation for the BDP 2.8 which gives the probability of observing a particular number of individuals as a Gaussian distribution with time dependent mean and variance (assuming the maximum population size N_i is sufficiently large (Ross, Pagendam & Pollett 2009)).

$$P_g(X = x_i; t) = \frac{1}{\sqrt{2\pi\sigma_i(t)^2}} exp\left(-\frac{x_i - n_i(t)}{2\sigma_i(t)^2}\right), \tag{2.12}$$

where the mean of this distribution ($n_i(t)$) is given by the solution of the logistic equation 3.6

$$n_i(t, n_{i0}) = \frac{K_i n_{i0} e^{r_i t}}{K_i + n_{i0}(e^{r_i t} - 1)}, \tag{2.13}$$

and the variance is given by

$$\sigma_i(t, n_{i0}) = N_i M_t^2 \int_0^t H_i(n_i(s, n_{i0})/N_i) M_s^{-2} ds, \tag{2.14}$$

where $M_s = \exp \int_0^s B_s ds$ and $B_s = F_i'(n_i(s)/N_i)$ (Ross, Taimre & Pollett 2006; Ross, Pagendam & Pollett 2009).

Bibliography

Allen, L. J. S. & Allen, E. J. (2003) A comparison of three different stochastic population models with regard to persistence time. *Theoretical Population Biology*, **64**, 439-449.

Alonso, D., McKane, A. J., & Pascual, M. (2007) Stochastic amplification in epidemics. *Journal of the Royal Society, Interface*, **4**, 575-582.

Beisner, B. E., Hovius, J., Hayward, A., Kolasa, J. & Romanuk, T. N. (2006) Environmental productivity and biodiversity effects on invertebrate community invasibility. *Biological Invasions*, **8**, 655-664.

Black, A. J., & McKane, A. J. (2012) Stochastic formulation of ecological models and their applications. *Trends in Ecology & Evolution*, **27**, 337-345.

Black, A. J., & McKane, A. J. (2010) Stochastic amplification in an epidemic model with seasonal forcing. *Journal of Theoretical Biology*, **267**, 85-94.

Blackburn, T. M., Cassey, P. & Lockwood, J. L. (2009) The role of species traits in the establishment success of exotic birds. *Global Change Biology*, **15**, 2852-2860.

Boys, R. J., Wilkinson, D. J., & Kirkwood, T. B. L. (2007) Bayesian inference for a discretely observed stochastic kinetic model. *Statistics and Computing*, **18**, 125-135.

Bright, C. (1998) *Life out of bound: bionvasions in a borderless world.* New York, NY: W.W: Norton & Company.

Cacho, J. O., Spring, D., Pheloung, P. & Hester, S. (2006) Evaluating the feasibility of eradicating an invasion. *Biological Invasions*, **8**, 903-917.

Cardinale, B. J., Srivastava, D. S., Duffy, J. E., Wright, J. P., Downing, A. L., Sankaran, M. & Jouseau, C. (2006) Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature*, **443**, 989-992.

Case, T. J. (1990) Invasion resistance arises in strongly communities model competition. *Proceedings of the National Academy of Sciences of the United States of America*, **87**, 9610-9614.

Cassey, P., Blackburn, T. M., Sol, D., Duncan, R. P., & Lockwood, J. L. (2004) Global patterns of introduction effort and establishment success in birds. *Proceedings of the Royal Society B: Biological Sciences*, **271**, 405-408.

Chapin, F. S., Zavaleta, E. S., Eviner, V. T., Naylor, R. L., Vitousek, P. M., Reynolds, H. L., Hooper, D. U., Lavorel, S., Sala, O. E., Hobbie, S. E., Mack, M. C & Díaz, S. (2000) Consequences of changing biodiversity. *Nature*, **405**, 234-242.

Chib, S., & Greenberg, E. (1995) Understanding the Metropolis-Hastings algorithm. *The American Statistician*, **49**, 327-335.

Corbin, J. D., & D'Antonio, C. M. (2004) Competition between native perennial and exotic annual grasses: implications for an historical invasion. *Ecology*, **85**, 1273-1283.

Drake, J. M., & Kramer, A. M. (2011) Mechanistic analogy: how microcosms explain nature. *Theoretical Ecology*, **5**, 433-444.

Drake, J. M., & Lodge, D. M. (2006) Allee effects, propagule pressure and the probability of establishment: risk analysis for biological invasions. *Biological Invasions*, **8**, 365-375.

Drake, J. M., Drury, K. L. S., Lodge, D. M., Blukacz, A., Yan, N. D., & Dwyer, G. (2006) Demographic stochasticity , environmental variability ,

and windows of invasion risk for Bythotrephes longimanus in North America. *Biological Invasions*, **8**, 843-861.

Drake, J. M. (2004) Allee effects and the risk of biological invasion. *Risk Analysis*, **24**, 795-802.

Drake, J. A., Huxel, G. R., & Hewitt, C. L. (1996) Microcosms as models for generating and testing community theory. *Ecology*, **77**, 670-677.

Ebenman, B., Law, R. & Borrvall, C. (2004) Community viability analysis: the response of ecological communities to species loss. *Ecology*, **85**, 2591-2600.

Elliott, E. C., & Cornell, S. J. (2013) Are anomalous invasion speeds robust to demographic stochasticity? *PloS One*, **8**, e67871.

Engen, S., Bakke, O., & Islam, A. (1998) Demographic and environmental stochasticity concepts and definitions. *Biometrics*, **54**, 840-846.

Filzbach, Microsoft Research Cambridge. `http://research.microsoft.com/en-us/um/cambridge/groups/science/tools/filzbach/filzbach.htm`

Galiana, N., Lurgi, M., Montoya, J. M., & López, B. C. (2014) Invasions cause biodiversity loss and community simplification in vertebrate food webs. *Oikos*, **123**, 721-728.

Gardinier, C. (2009) *Stochastic methods: a handbook for the natural and social sciences*. Springer.

Gilioli, G., Pasquali, S., & Ruggeri, F. (2008) Bayesian inference for functional response in a stochastic predator-prey system. *Bulletin of Mathematical Biology*, **70**, 358-381.

Gillespie, D. T. (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, **22**, 403-434.

Gillespie, D. T. (1977) Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, **81**, 2340-2361.

Giometto, A., Rinaldo, A., Carrara, F. & Altermatt, F. (2014) Emerging predictable features of replicated biological invasion fronts. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 297-301.

Hastings, A., Cuddington, K., Davies, K. F., Dugaw, C. J., Elmendorf, S., Freestone, A., Harrison, S., Holland, M., Lambrinos, J., Malvadkar, U., Melbourne, B. A., Moore, K., Taylor, C. & Thomson, D. (2004) The spatial spread of invasions: new developments in theory and evidence. *Ecology Letters*, **8**, 91-101.

Jeschke, J. M. & Strayer, D. L. (2006) Determinants of vertebrate invasion success in Europe and North America. *Global Change Biology*, **12**, 1608-1619.

Keane, R. M., & Crawley, M. J. (2002) Exotic plant invasions and the enemy release hypothesis. *Trends in Ecology & Evolution*, **17**, 164-170.

Kenis, M., Auger-Rozenberg, M. A., Roques, A., Timms, L., Péré, C., Cock, M. J. W., Settele, J., Augustin, S. & Lopez-Vaamonde, C. (2008) Ecological effects of invasive alien insects. *Biological Invasions*, **11**, 21-45.

Kimura, M. (1964) Diffusion models in population genetics. *Journal of Applied Probability*, **2**, 177-232.

Kolar, C. S., & Lodge, D. M. (2002) Ecological predictions and risk assessment for alien fishes in North America. *Science*, **298**, 1233-1236.

Lande, R., Enger, S. & Sæther, B. E. (2009) *Stochastic population dynamics in ecology and conservation.* Oxford Series in Ecology and Evolution, Oxford

Leslie, P. H. (1957) An analysis of the data for some experiments carried out by Gause with populations of the protozoa, *paramecium aurelia* and *paramecium caudatum*. *Biometrika*, **44**, 314-327.

Litchman, E. (2010) Invisible invaders: non-pathogenic invasive microbes in aquatic and terrestrial ecosystems. *Ecology Letters*, **13**, 1560-1572.

Lotka, A. J. (1920) Analytical note on certain rhythmic relations in organic systems. *Proceedings of the National Academy of Sciences of the United States of America*, **6**, 410-415.

Lurgi, M., Galiana, N., López, B. C., Joppa, L. & Montoya, J. M. (2014) Network complexity and species traits mediate the effects of biological invasions on dynamic food webs. *Manuscript*

McKane, A. J. & Newman, T. J. (2004) Stochastic models in population biology and their deterministic analogs. *Physical Review E*, **70**, 1-19.

McKane, A. J. & Newman, T. J. (2005) Predator-prey cycles from resonant amplification of demographic stochasticity. *Physical Review Letters*, **94**, 1-4.

Nåsell, I. (2001) Extinction and quasi-stationarity in the Verhulst logistic model. *Journal of Theoretical Biology*, **211**, 11-27.

Nisbet, R. M. & Gurney, W. S. C. (1982) *Modelling fluctuating populations* Blackburn Press, New Jersey.

Ovaskainen, O., & Meerson, B. (2010) Stochastic models of population extinction. *Trends in Ecology & Evolution*, **25**, 643-652.

Palamara, G. M., Delius, G. W., Smith, M. J., & Petchey, O. L. (2013) Predation effects on mean time to extinction under demographic stochasticity. *Journal of Theoretical Biology*,

Palamara, G. M., Childs, D. Z., Clements, C., Plebani, M., Petchey, O. L. & Smith, M. (2014) Inferring the temperature dependence of ecological parameters: the effect of the experimental design and the inference method. *Ecology & Evolution*, In Press.

Pascual, M. A., & Kareiva, P. (1996) Predicting the outcome of competition using experimental data: maximum likelihood and Bayesian approaches. *Ecology*, **77**, 337-349.

Pereira, H. M., Leadley, P. W., Proença, V., Alkemade, R., Scharlemann, J. P. W., Fernandez-Manjarrés, J. F., Araújo, M. B., Balvanera, P., Biggs, R., Cheung, W. W. L., Chini, L., Cooper, H. D., Gilman, E. L., Guénette, S., Hurtt, G. C., Huntington, H. P., Mace, G. M., Oberdorff, T., Revenga, C., Rodrigues, P., Scholes, R. J., Sumaila, U. R. & Walpole, M. (2010) Scenarios for global biodiversity in the 21st century. *Science*, **330**, 1496-1501.

Pollett, P. K. (1992) Diffusion Approximations for Ecological Models.

Romanuk, T. N., Zhou, Y., Brose, U., Berlow, E. L., Williams, R. J. & Martinez, N. D. (2009) Predicting invasion success in complex ecological networks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **364**, 1743-1754.

Ross, J. V., Pagendam, D. E. & Pollett, P. K. (2009) On parameter estimation in population models II: multi-dimensional processes and transient dynamics. *Theoretical Population Biology*, **75**, 123-132.

Ross, J. V., Taimre, T. & Pollett, P. K. (2006) On parameter estimation in population models. *Theoretical Population Biology*, **70**, 498-510.

Sax, D. F., Stachowicz, J. J., Brown, J. H., Bruno, J. F., Dawson, M. N., Gaines, S. D., Grosberg, R. K., Hastings, A., Holt, R. D., Mayfield, M. M., O'Connor, M. I. & Rice, W. R. (2007) Ecological and evolutionary insights from species invasions. *Trends in Ecology & Evolution*, **22**, 465-471.

Shea, K., & Chesson, P. (2002) Community ecology theory as a framework for biological invasions. *Trends in Ecology & Evolution*, **17**, 170-176.

Simberloff, D., Martin, J. L., Genovesi, P., Maris, V., Wardle, D. A., Aronson, J., Courchamp, F., Galil, B., García-Berthou, E., Pascal, M., Pyšek,

P., Sousa, R., Tabacchi, E. & Vilá, M. (2013) Impacts of biological invasions: what's what and the way forward. *Trends in Ecology & Evolution*, **28**), 58-66.

Simberloff, D. (2011) Charles Elton: neither founder nor siren, but prophet. In *Fifty Years of Invasion Ecology: The Legacy of Charles Elton* Richardson, D.M., ed., pp. 11-24, Blackwell Publishing.

Snyder, R. E. (2003) How demographic stochasticity can slow biological invasions. *Ecology*, **84**, 1333-1339.

Taylor, C. M., & Hastings, A. (2005) Allee effects in biological invasions. *Ecology Letters*, **8**, 895-908.

Tilman, D. (2004) Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences of the United States of America*, **30**, 10854-10861.

Toni, T., Welch, D., Strelkowa, N., Ipsen, A. & Stumpf, M. P. H. (2009) Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of The Royal Society Interface*, **6**, 187-202.

Van Kleunen, M., Dawson, W., Schlaepfer, D., Jeschke, J. M. & Fischer, M. (2010) Are invaders different? A conceptual framework of comparative approaches for assessing determinants of invasiveness. *Ecology Letters*, **13**, 947-958.

Vilá, M., Espinar, J. L., Hejda, M., Hulme, P. E., JarošÃk, V., Maron, J. L., Pergl, J., Schaffner, U., Sun, Y. & Pyšek, P. (2011) Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. *Ecology Letters*, **14**, 702-708.

Vitousek, P. M., Antonio, C. M. D., Loope, L. L. & Westbrooks, R. (1996) Biological invasions as global environmental change. *American Scientist*, **84**, 468-478.

Wilkinson, D. (2006) *Stochastic modelling for systems biology.* Chapman & Hall/CRC.

Williams, R. J. (2008). Effects of network and dynamical model structure on species persistence in large model food webs. *Theoretical Ecology*, **1**, 141-151.

Williamson, M. H. & Fitter, A. (1996) The characters of successful invaders. *Biological Conservation*, **78**, 163-170.

With, K. A. (2002). The landscape ecology of invasive spread. *Conservation Biology*, **16**, 1192-1203.

Yodzis, P., & Innes, S. (1992) Body size and consumer-resource dynamics. *The American Naturalist*, **139**, 1151-1175.

Chapter 3

Predation effect on mean time to extinction under demographic stochasticity

Gian Marco Palamara, Gustav W. Delius, Matthew J. Smith & Owen L. Petchey

Journal of Theoretical Biology, **334**, 61-70. (2013)

3.1 Abstract

Methods for predicting the probability and timing of a species' extinction are typically based on single species population dynamics. Assessments of extinction risk often lack effects of interspecific interactions. We study a birth and death process in which the death rate includes the effect of predation. Predation is included via a general nonlinear expression for the functional response of predation to prey density. We investigate the effects of the foraging parameters (e.g. attack rate and handling time) on the mean time to extinction. Mean time to extinction varies by orders of magnitude when we alter the foraging parameters, even when we exclude the effects of these parameters on the equilibrium population size. Conclusions are robust to assumptions about initial conditions and variable predator abundance. These findings clearly show that accounting for the nature of interspecific interactions is likely to be critically important when estimating extinction risk.

Key Words: Trophic Interaction, Predator-Prey Model, Birth and Death Process, Quasistationary Distribution.

3.2 Introduction

Population ecologists have long sought to understand how exogenous factors (such as environmental variability) interact with endogenous factors (such as body size, life history, trophic position) to determine the probability that a population will go extinct (Lande, Steinar & Sæther 2004). Such insights promise improvements in our understanding of species population dynamics in time and space (Bascompte & Solé 1995), in the determinants of ecosystem stability and complexity (McCann 2000; Allesina & Tang 2012), and practical advances in our ability to conserve or eradicate populations (Witting, Tomiuk & Loeschcke 2000; Liebhold & Bascompte 2003).

Populations can take different routes to extinction. Extinction can occur either through progressive declines in population size, such as due to habi-

tat deterioration, or through sudden crashes in abundance, such as through random catastrophes (Lande 1993). Demographic stochasticity is caused by random variation among individuals in survival and reproduction. Environmental stochasticity is, on the other hand, random variation in the environment, which can lead to changes in the rates of processes influencing population dynamics (such as survival and reproduction rates; Lande, Steinar & Sæther (2004)). While environmental stochasticity can be important for populations of any size, demographic stochasticity becomes particularly important at low population sizes.

Classical population theory shows that if demographic stochasticity is the only stochastic process influencing the dynamics of a population then the mean time to extinction increases exponentially with equilibrium population size. In contrast, environmental stochasticity alone can lead to a power law relationship between the mean time to extinction and equilibrium population size (Lande 1993; Foley 1994). A major body of research to date has sought to understand how the characteristics of individual species influence their probability of extinction. For example, slow life histories and small geographical range sizes are associated with a high extinction risk (Purvis *et al.* 2000). One outcome of this large body of research is population viability analysis (PVA; Brook *et al.* (2000)). PVA combines the effects of these different factors to estimate the overall probability that a population will go extinct (Beissinger & McCullough 2002; Mace *et al.* 2008). However the single species models used for assessing population viability often lack the explicit incorporation of direct trophic interactions (Sabo & Gerber 2007; Sabo 2008).

Interspecific trophic interactions have been widely studied, theoretically and experimentally, in the fields of population and community ecology. In his pioneering work Holling (Holling 1959) proposed a simple non-linear relationship between prey density and predator feeding rate, known as the *predator functional response*, that is still widely used today. Since then, various modifications to Holling's original formulation have been made to represent different

foraging mechanisms (Real 1977; Abrams & Ginzburg 2000; Jeschke, Kopp & Tollrian 2002). A general expression for these is

$$f(n) = \frac{\alpha n^q}{1 + h\alpha n^{q+1}}, \tag{3.1}$$

where αn^q is the attack rate (a measure of the encounter rate and capture success of the predator foraging on the prey), h is the handling time (a measure of the time needed to attack, eat and digest the prey). The attack rate term (αn^q) allows for the different scalings of attack rate with prey density that, in combination with the handling time, define the three standard types of functional response:

- Setting $q = 0$ and $h = 0$ in expression (3.1) we obtain the type I functional response. In this case predation rate increases linearly with prey abundance and the predator has a negligible handling time or is able to search and capture prey while handling other prey (Holling 1965; Jeschke, Kopp & Tollrian 2004). Type I functional responses are classically associated with filter feeders (Jeschke, Kopp & Tollrian 2004).

- Setting $q = 0$ in expression (3.1) we obtain the type II functional response (Holling 1959; Real 1977). The type II functional response is the simplest expression that takes into account the time taken for predators to locate and consume ("handle") their prey. The type II functional response is classically associated with specialist predators (Turchin 2003).

- When $h > 0$ and $q > 0$ we obtain the type III functional response. The type III has been associated with learning effects of the predator in catching and handling its prey (Real 1977), with generalist predators switching among alternate prey (Smout *et al.* 2010) or with spatial effects enabling prey to hide from predators (Vucic-Pestic *et al.* 2010). Both the type II and III functional responses are characterized by a maximum intake rate ($1/h$) at which predation rate saturates (Holling 1959)

The effect of the predator functional response on population stability has been widely studied (May 2001; Drossel, McKane & Quince 2004). However, such studies have predominantly been conducted using deterministic models (they have also addressed various stability concepts). Nevertheless interspecific interactions are stochastic events. Therefore characteristics of their stochasticity are also likely to influence extinction time. For example, predation pressure could reduce population size and thereby increase the chance of extinction via demographic stochasticity. We should also expect that different predator foraging behaviors will have different effects on the population's risk of extinction. Most PVAs incorporate the effects of interspecific interactions in population level parameters (Sabo & Gerber 2007), for example, predation is included as a constant (density independent) source of mortality rather than a coupled, density dependent population process (Sabo 2008). Single species models used in PVA typically fail to incorporate dynamic interactions between populations and their predators, which can bias population viability estimates (Sabo & Gerber 2007; Sabo 2008).

The stochastic effects of interspecific interactions have also been investigated in multispecies models. Stochastic models of population dynamics range from predator-prey models (McKane, & Newman 2004) to stochastic food web models (Ebenman, Law & Borrvall 2004; Powell & Roland 2009). These models represent trophic-interactions as individual-based reaction probabilities. In cases with large numbers of individuals, these models produce macroscopic properties that can be predicted by their deterministic counterparts (McKane, & Newman 2004; Black & McKane 2012). However, the stochastic models can also exhibit properties that cannot be predicted by the deterministic models, for example, stochastic resonance (McKane, & Newman 2005). Extinction rates have also been investigated in predator prey models with linear interaction rates (Parker & Kamenev 2009). The effects of non linear birth rates on extinction rates have also been studied (Liebhold & Bascompte 2003; Kramer & Drake 2010). However, to our knowledge, little is known about the effects of the foraging parameters of the functional response on extinction rates.

In this paper we investigate how predation influences the population dynamics of small populations experiencing demographic stochasticity. We aim to clarify the effects of predation on the extinction process. We analyze a single species birth-death process in which the death rate includes density dependent predation by a predator whose abundance is not affected by the abundance of its prey (we explain this assumption below). We then investigate the importance of the predation functional response parameters on the mean time to extinction and examine the robustness of our results to temporal variation in predator abundance. This investigation is novel in itself and represents an initial step towards a complete appreciation of the effects of interspecific interactions on time to extinction.

3.3 Methods

We define a single species population as having n individuals at time t. We next define functions for the birth rate $b(n)$ and death rate $d(n)$ of individuals in the population as

$$\begin{aligned} b(n) &= \lambda\left(1-\frac{n}{k}\right), \\ d(n) &= \mu + f(n)z, \end{aligned} \tag{3.2}$$

where z is the abundance of predators. The logistic form of the birth rate function represents density dependent effects such as intraspecific competition for resources. The parameter λ is the per capita birth rate of the prey species in the absence of density dependence and k is the maximum possible population size (Nisbet & Gurney 1982; Nåsell 2001). The death rate includes a constant term μ, the predator-free per capita death rate, and the term $f(n)z$, the per capita functional response of predation rate to prey abundance (expression (3.1)). The expected maximum per capita growth rate of the population is then the difference between per capita birth and death rates i.e., $r = \lambda - \mu$. We will keep the per capita birth rate and the predator-free per capita death rate fixed and change the parameters of the

functional response in an interval constrained by biological arguments:

- Handling time h takes values between $0.001d$ and $0.1d$[1]. We assume that h cannot be larger than the average lifetime of the prey species (assumed to be $1/\lambda$).

- Attack rate (αn^q) takes values between $0.001d^{-1}$ and $10d^{-1}$. This choice is justified by empirical observations of foraging behavior of insects and other organisms (Vucic-Pestic *et al.* 2010; Hammill, Petchey & Anholt 2010; Hanski, Hansson & Henttonen 1991).

- The exponent q takes values between 0 and 2 again due to empirical observations (Vucic-Pestic *et al.* 2010) and theoretical work (Williams & Matinez 2004).

- Predator abundance z is kept independent from the prey as we aim to study the effect of a generalist predator. In most of our analyses below we fix z to be constant, the simplest manifestation of the assumption that predator abundance is independent of prey abundance, as has been assumed for generalist predators in previous studies (e.g. Hanski, Hansson & Henttonen (1991)). In this case, when the growth parameters λ, μ and k are fixed, the relevant parameters for the dynamics are αz and h/z (see appendix 3.6.1). Later, we relax the assumption that z is constant and allow it to fluctuate (details below). We choose z to take values between 1 and k.

We combine expressions (3.1) and (3.2) to obtain expressions for the population birth and death rates, $B(n)$ and $D(n)$, respectively as

$$\begin{aligned} B(n) &= nb(n) = n\lambda\left(1-\frac{n}{k}\right), \\ D(n) &= nd(n) = n\mu + \frac{\alpha z n^{q+1}}{1+\alpha h n^{q+1}}. \end{aligned} \tag{3.3}$$

The state of the system can be characterized by the probability $p(n,t)$ of having n individuals at time t, where n takes integer values in the range

[1]The unit of measure of time is arbitrarily set to days (d).

$\{0, \ldots, k\}$. The master equation describing the time evolution of the probability distribution is

$$\begin{aligned}\frac{dp(n,t)}{dt} =& D(n+1)p(n+1,t) + B(n-1)p(n-1,t) \\ & - (B(n) + D(n))p(n,t).\end{aligned} \tag{3.4}$$

In order to express equation (3.4) in a compact way we defined $p(k+1,t) = p(-1,t) = 0$. Given an initial probability distribution $p(n,0)$, equation (3.4) uniquely determines the probability distribution at later times. The process represented in this master equation always ends up with population extinction, represented as the stationary distribution $\overline{\mathbf{p}} = (1, 0, \ldots, 0)$, which the distribution $p(n,t)$ approaches as time t approaches infinity. This is because, given sufficient time (which may be extremely large when populations are large) stochastic fluctuations in the population size will eventually, at some time, cause the population size to be zero. Since there is no immigration in this model the population cannot recover from that state. In section 3.3.2 we will show how to compute the probability of having a population of size n at time t conditioned on the fact that it has not yet gone extinct.

3.3.1 Deterministic rate equation

A deterministic rate equation, describing the time evolution of the mean population size, can be associated to every stochastic birth and death process (Gardinier 2009). The deterministic equation is given by

$$\frac{dn}{dt} = B(n) - D(n). \tag{3.5}$$

In our case, substituting the rates (3.3) into equation (3.5), we obtain

$$\frac{dn}{dt} = \lambda n \left(1 - \frac{n}{k}\right) - \mu n - \frac{\alpha z n^{q+1}}{1 + \alpha h n^{q+1}}. \tag{3.6}$$

In order to understand the roles of deterministic versus stochastic effects in prey extinction, we compute the fixed points of equation (3.6) and their stability. We use as measure of stability the real part of the leading eigenvalue

of the Jacobian matrix evaluated at each fixed point (hereafter termed the stability coefficient, detailed in appendix 3.6.1). Mean population numbers close to the value of an unstable fixed point evolve away from the fixed point value while those close to a stable fixed point value evolve towards that value.

3.3.2 Deterministic and stochastic effects on extinction

In equation (3.6) the extinction state $n = 0$ is always a fixed point, while the other fixed points are given by the intercepts between the per capita growth term $(r - \lambda n/k)$ and the per capita functional response $zf(n)$ (see figure 3.2). Two different extinction scenarios are revealed by the analysis of the deterministic equation, associated with changes in the stability of the extinction state (Assaf & Meerson 2010).

When $q = 0$ and $h > 0$ (type II functional response) the extinction state can be stable or unstable. A stable extinction state indicates that when the population size is sufficiently close to zero the birth-death process (3.3) causes the mean population size to tend towards zero because the mean population death rate exceeds the mean birth rate. In contrast, an unstable extinction state indicates that when the mean population size is sufficiency close to zero the population size tends away from zero, and demographic stochasticity is necessary for the population to go extinct. If $\alpha_p = z\alpha > r$, the extinction state is stable. In this case, there are two positive fixed points, one stable and one unstable, or no positive fixed points (see figure 3.1). At $\alpha_p = r$ there is a transcritical bifurcation, at which the unstable fixed point goes to 0 (see figure 3.1) and the extinction state becomes unstable. When the extinction state is unstable ($\alpha_p < r$) there is only one other fixed point and this is stable (see figure 3.1).

When $q > 0$ and $h > 0$ (type III functional response), the model has at most three non zero fixed points and the extinction state will always be unstable. In contrast to the type II functional response, for $q > 0$ there are combi-

nations of foraging parameters for which there can be up to three non zero fixed points (see appendix 3.6.1). In this case, two of the three non-zero fixed points are stable, and one is unstable (see figure 3.1).

Given the birth and death process (3.3), extinction is caused by demographic stochasticity and typically occurs in two different ways depending on whether the extinction state is a stable or unstable fixed point of the deterministic rate equation (3.6). For the type II functional response, below the transcritical bifurcation, extinction is caused by a large fluctuation which brings the system from the stable to below the unstable fixed point. From there a fast deterministic evolution takes the population quickly to the stable extinction state. For the type II functional response after the bifurcation, and for the type III functional response, extinction is caused by a large fluctuation in density which brings the system from the stable fixed point to the unstable, absorbing extinction state. For the type III functional response when there are two stable fixed points and the population is near the larger stable fixed point, extinction needs two large fluctuations to occur, one fluctuation which brings the system from the larger to the lower stable fixed point and another fluctuation which brings the system from the lower stable fixed point to the unstable extinction state.

3.3.3 Quasistationary distribution (QSD) and mean time to extinction (MTE)

When the deterministic rate equation has at least one stable fixed point, the system approaches a quasi stationary state with a time independent distribution $\pi(n)$; this is called the *Quasistationary Distribution* (Bartlett 1960; Nisbet & Gurney 1982). The quasistationary distribution $\pi(n)$ is obtained from the probability $p_c(n,t)$, that of finding n individuals at time t, conditioned on the fact extinction has not occurred yet:

$$p_c(n,t) = \frac{p(n,t)}{1-p(0,t)}. \tag{3.7}$$

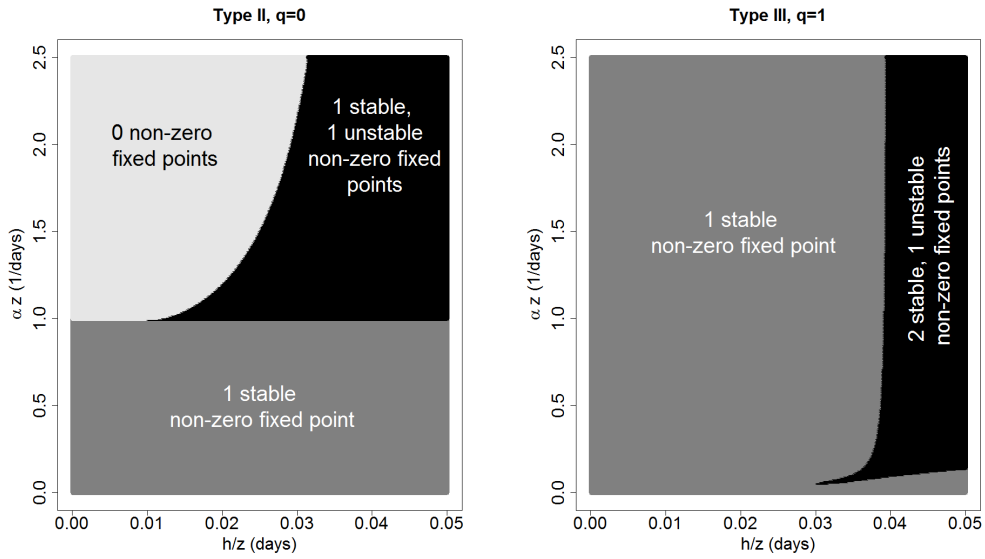

Figure 3.1: Regions in parameter space where the model (3.6) has 0, 1, 2 and 3 non-zero fixed points and their associated stability. The axes are handling time divided by predator abundance h/z and attack rate multiplied by predator abundance $\alpha_p = z\alpha$. The growth parameters are fixed as $\lambda = 1.5\,d^{-1}$; $\mu = 0.5\,d^{-1}$ and $k = 150$.

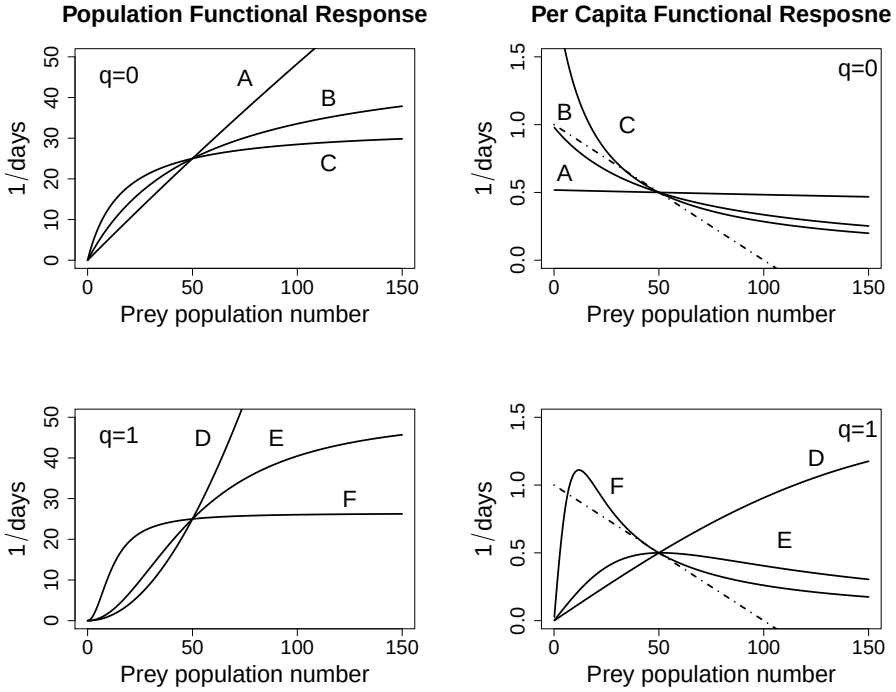

Figure 3.2: Prey population mortality due to predation (left panels) and per capita mortality due to predation (right panels) following a type II (upper panels) and a type III (lower panels) functional response. On the left plots solid lines are the population functional responses. On the right plots solid lines are the per capita functional responses while the dashed line is the per capita growth curve ($\lambda = 1.5\,d^{-1}$; $\mu = 0.5\,d^{-1}$ and $k = 150$). The foraging parameters are for the type II functional response A: $h/z = 0.001\,d$; $\alpha z = 0.5\,d^{-1}$. B: $h/z = 0.02\,d$; $\alpha z = 1\,d^{-1}$ at the transcritical bifurcation. C: $h/z = 0.03\,d$; $\alpha z = 2.05\,d^{-1}$. And for the type III functional response D: $h/z = 0.001\,d$; $\alpha z = 0.01\,d^{-1}$. E: $h/z = 0.02\,d$; $\alpha z = 0.02\,d^{-1}$. F: $h/z = 0.18\,d$; $\alpha z = 0.037\,d^{-1}$. Note that with this particular choice of foraging parameters the value of the fixed point is fixed at $n_0 = 50$.

We derive a master equation for the conditioned probability $p_c(n,t)$ and look for its stationary solution $\pi(n)$ (see appendix 3.6.2). When the initial condition (the probability $p_0(n)$ of finding n individuals at time 0) of equation (3.4) is set to the quasistationary distribution, then the probability of finding n individuals at time t becomes

$$p(n,t) \simeq \pi(n) \exp(-t/MTE). \tag{3.8}$$

The time to extinction is then an exponentially distributed random variable with mean equal to the MTE and,

$$MTE = \frac{1}{D(1)\pi(1)}. \tag{3.9}$$

There are more complicated expressions for the mean time to extinction when the initial condition is not the quasi stationary distribution $\pi(n)$ (see appendix 3.6.2). We found the time for the system to reach the quasistationary distribution is negligibly small compared to the MTE. We compute the MTE of the birth and death process (3.3) for different functional responses and different values of the foraging parameters within the functional responses. All other parameters remain fixed.

3.3.4 Numerical calculations

It is not possible to obtain closed expressions for the quasistationary distribution $\pi(n)$ of the birth and death process (3.3). Instead, we obtain the quasistationary distribution in a realistic range of foraging parameters by an iterative numerical scheme described in appendix 3.6.2. In order to perform numerical calculations we fix the growth parameters of equation (3.6) in the following way: $\lambda = 1.5\, d^{-1}$; $\mu = 0.5\, d^{-1}$; $k = 150$. With that choice the intrinsic growth rate of the prey population is fixed to $1\, d^{-1}$. In the absence of predators i.e., setting $z = 0$ in equation (3.6), the model has a non zero fixed point at $n_0 = kr/\lambda = 100$.

We compute the logarithm of the MTE of the birth and death process (3.3)

obtained using the quasi stationary distribution as initial condition (see appendix 3.6.2) for different combinations of foraging parameters. The foraging parameters of the functional response affect the MTE in direct and indirect ways. The foraging parameters affect the MTE indirectly by influencing the equilibrium population size of the prey. They influence it directly by changing the MTE even for a fixed equilibrium prey population size. In order to isolate the effect of the shape of the functional response (direct effect) from the effect of the fixed point population number n_0 (indirect effect), we keep n_0 unchanged. This imposes the following relation between the attack rate parameter α and the handling time parameter h:

$$\alpha_q = \frac{\alpha(h, n_0)}{n_0^q} = \frac{rk - \lambda n_0}{n_0^q[zk - hn_0(rk - \lambda n_0)]}. \tag{3.10}$$

In order to avoid fixing an unstable equilibrium with relation (3.10), we limit our investigation to those values of handling time that give rise to a stable fixed point (see appendix 3.6.1) i.e.,

$$h < h_1 = \frac{\lambda zk[(1-q)\lambda n_0 + qrk]}{n_0(1+q)(rk - \lambda n_0)^2}. \tag{3.11}$$

As the handling time approaches the value

$$h_0 = \frac{zk}{n_0(rk - \lambda n_0)}, \tag{3.12}$$

the required attack rate approaches infinity, so we limit ourselves to $h < \min(h_0, h_1)$. In the case of $q > 0$ we also keep h small enough so that we do not enter the bistable region. Matlab code to reproduce our calculations can be found at `http://purl.org/net/extinction_code`.

3.3.5 Temporal variation in predator abundance

Finally, we explored the effects of variation in predator abundance on the MTE. Here we do not have a closed expression for the MTE of the prey but we can simulate many replicates of the birth and death process (3.3) using

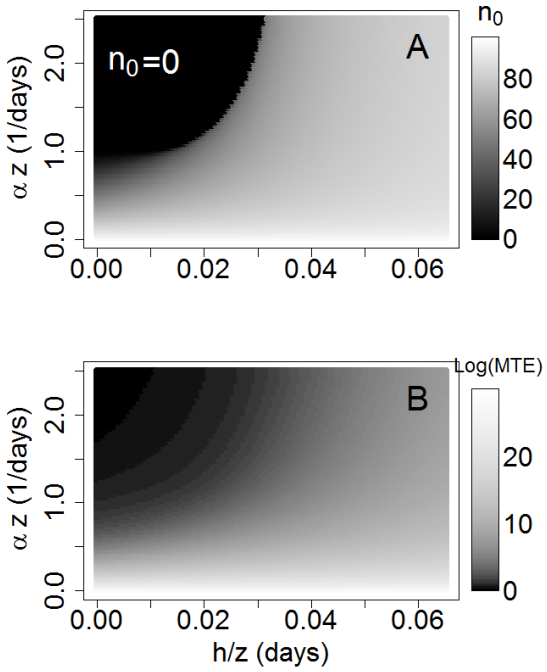

Figure 3.3: Abundance n_0 at the stable fixed point predicted by equation (3.6) (A) and the log of the mean time to extinction (MTE) (B) of the birth and death process (3.3) using the quasistationary distribution as initial condition as a function of handling time and attack rate for type II functional response. The growth parameters are as specified in the legend of figure 3.1.

the Gillespie algorithm (Gillespie 1977) and compute the MTE numerically. The assumption of constant predator abundance was relaxed by allowing z (predator abundance) to vary randomly through time, with values drawn from a uniform distribution $[z - z/2, z + z/2]$, where separate investigations were performed for $z = 2, 10, 50$. We then explored the effects of seasonal variation in predator abundance on the MTE of the prey. Periodic forcing was introduced by setting

$$z = z_0(1 - \sin(\omega t)), \tag{3.13}$$

where z_0 is the amplitude and ω is the frequency of the forcing. We performed separate investigations for $z0 = 2, 10, 50$ and with $\omega = 0.1, 0.02, 0.01$.

3.4 Results

The MTE of the model without predation is extremely large ($10^{25}d$), meaning that extinction will almost certainly never occur. When there are predators, changes in the attack rate (α), the handling time (h) and the scaling of attack rate (q) strongly influences the MTE through changes in equilibrium population size. These indirect effects on the MTE can be seen in figure 3.3. As expected, the MTE is relatively low when the functional response parameters lead to deterministic extinction (region labeled $n_0 = 0$ in figure 3.3). When the foraging parameters are such that there is a stable positive equilibrium, the MTE (figure 3.3 B) is, unsurprisingly, positively related to the equilibrium population size (figure 3.3 A). In the region of the parameter space where there are three non-zero fixed points (type III functional response, figure 3.1 right panel) we found extinction occurring at evolutionary time scales e.g. MTE $> 10^{15}$. Therefore, we did not perform a detailed investigation of the model in this region of the parameter space.

Figures 3.4 and 3.5 show the direct effect of the foraging parameters on the MTE and stability coefficient (i.e., the effect when the equilibrium density is kept fixed). These show that, when the extinction state is unstable, the MTE decreases less than exponentially with handling time (figure 3.4), or, in the case of type III functional response, more than exponentially (figure 3.5). Also when the extinction state is stable, the MTE decreases less than exponentially (figure 3.4). We fitted exponential curves to the MTE for three different values of n_0, with type II functional response, before the transcritical bifurcation i.e. when the extinction state is unstable. We observe an increase in both the slope and the intercept with increasing carrying capacity[2].

The stability of the fixed point (for details see appendix 3.6.1) also has an effect on the MTE. From the lower panels of figures 3.4 and 3.5 we see that

[2]The values of the slope are fitted using a least square method and are, for different values of the fixed point: $n_0 = 50$, slope -133.7; $n_0 = 60$, slope -168.3; $n_0 = 70$ slope -173.

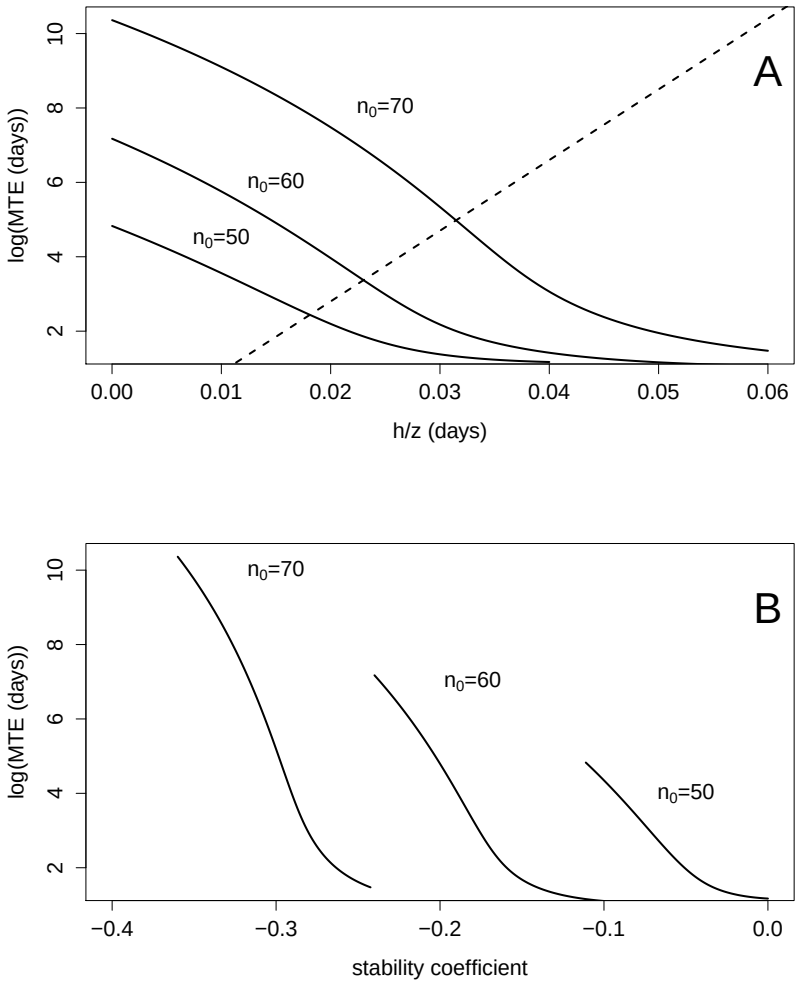

Figure 3.4: Logarithm of the mean time to extinction (MTE) (when the quasistationary distribution is set as initial condition of equation (3.4)) as a function of handling time (A) and of the stability coefficient (B), for different values of the stable fixed point n_0, for a type II functional response. The dashed line divides the parameter space into two regions. A region associated with an unstable extinction state (above and left) and a region associated with a stable extinction state (below and right). The curves are drawn for all values of handling time which keep the stable equilibrium density n_0 fixed and the corresponding attack rate (3.10) finite i.e., for $h < min(h_1, h_0)$. The values of h_1 and h_0 are obtained from expressions (3.11) and (3.12). We used expression (3.9) to obtain the MTE. The stability coefficient is the real part of the Jacobian computed in n_0 (see appendix 3.6.1). The growth parameters are as specified in the legend of figure 3.1.

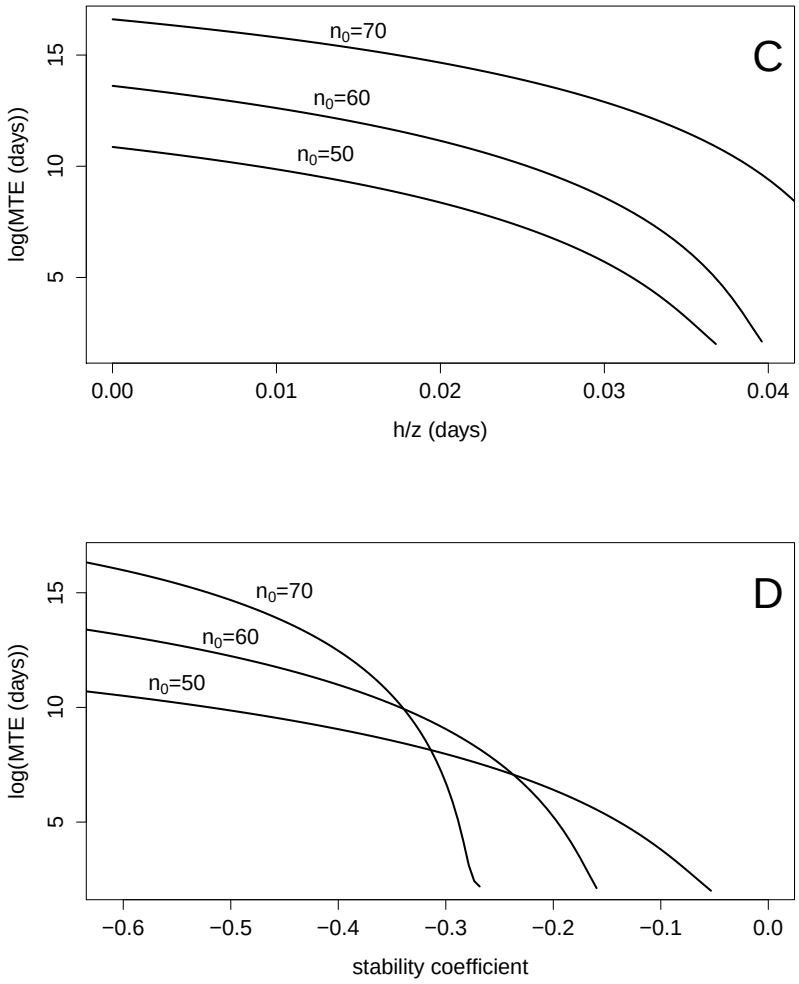

Figure 3.5: Logarithm of the mean time to extinction (MTE) as a function of handling time (C) and of the stability coefficient (D), for different values of the stable fixed point n_0, for type III functional response ($q = 1$). We used expression (3.9) to obtain the MTE. The growth parameters are as specified in the legend of figure 3.1.

type II functional response gives raise to less stable fixed points than type III i.e., to a less negative stability coefficient. Consequently, at the same value of the stable fixed point, the MTE associated with the type II functional response is lower than the MTE associated with the type III. Moreover, while keeping the stable equilibrium fixed, the variation in MTE with handling time spans a larger interval of stability coefficient for the type III than for the type II functional response. These differences in the range of stability produce the different declines in MTE with handling time for the different functional responses (steeper for the type II figure 3.4 upper panel).

Overall, when we keep the stable equilibrium fixed we observe variation in the MTE by up to 5 orders of magnitude for a type II functional response and by up to 10 orders of magnitude when we have a type III functional response. This variation in MTE is shown as a function of handling time and predator abundance in figures 3.4 and 3.5 when the growth parameters are fixed and attack rate varies with handling time according to equation (3.10).

In figure 3.6 we show the effects of handling time on the QSD while keeping the stable equilibrium fixed. The width of the QSD increases with increasing handling time. Since equilibrium population size is kept constant, this increasing handling time corresponds to an increasing attack rate (expression (3.10)), and these both make extinction more likely (reducing the MTE). It is possible to see the influence of the unstable fixed point in the left tail of the QSD for the type II functional response (figure 3.6 left panels). Below the transcritical bifurcation (figure 3.6 A) the extinction state is unstable. After the bifurcation (figure 3.6 B and C) the extinction state becomes stable due to the emergence of a positive unstable fixed point, making the left tail of the QSD increase as population size decreases towards zero. For this reason, in the parameter region after the transcritical bifurcation, the MTE of the model (3.4) becomes dependent on the initial population distribution. In figure 3.7 we show how the effect of using a delta function instead of the QSD as initial distribution can decrease the MTE by up to 2 orders of magnitude, if the initial condition is close to the (stable) extinction state (Detailed

methods used to compute the MTE with different initial distributions can be found in appendix 3.6.2).

Lastly, we show how allowing predator abundance to vary interacts with the effects of the foraging parameters on the MTE for both type II and type III functional responses. When predator abundance z fluctuates randomly through time, the average MTE is unchanged. Consequently, the effect of foraging parameters (h/z) on the MTE remains unchanged. Periodic variation in predator abundance can, on average, reduce or increase the MTE depending on the value of the foraging parameters (figure 3.8). The nature of this effect is such that the MTE decreases at low handling time and increases at high handling times. This tends to cause a shallower negative relationship between the MTE and the foraging parameters. Moreover we note that, for the type II functional response, an increase in the frequency of the periodic variation in predator abundance (ω) causes a larger decrease in the MTE (figure 3.8 A); while, for the type III functional response, ω has no effect on the variation in MTE (figure 3.8 B).

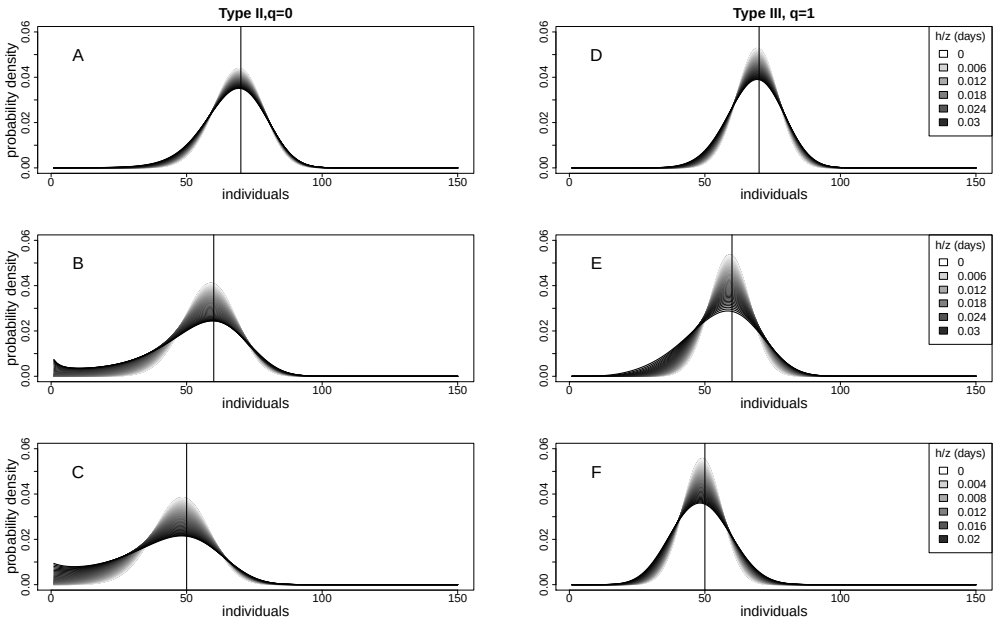

Figure 3.6: Quasistationary distribution of the model (3.4) when keeping the stable equilibrium fixed at $n_0 = 70$ (A, D), $n_0 = 60$ (B, E), $n_0 = 50$ (C, F) for type II (left panels) and type III (right panels) functional response. The gray scale represents increasing handling time over fixed predator abundance.

3.5 Discussion

We have shown how different choices of the foraging parameters vary the mean time to extinction by up to 10 orders of magnitude even when equilibrium population size is kept constant. Therefore, our results imply that estimates of extinction risk could be extremely inaccurate without explicitly accounting for interspecific interactions.

There is a wide literature describing experimental measures of foraging parameters (attack rate, handling time and scaling exponent) (Rall *et al.* 2012). These studies include predator-prey interactions among terrestrial and aquatic organisms such as protists (Hammill, Petchey & Anholt 2010) and arthropods (Spitze 1985; Smout *et al.* 2010). However, similar measurements of the nature of predator-prey interactions are absent in most of the studies related to the extinction risk of individual species (Sabo 2008). We found that accounting for different foraging strategies (i.e. different functional responses and different foraging parameters) can become critical when evaluating the extinction risk of a target prey species (Prowse *et al.* 2013).

The effects of the foraging parameters can be intuitively explained by examining the strength of population regulation at the equilibrium density i.e., the slope of the functional response at the fixed point (Figure 3.2). Keeping the stable equilibrium fixed requires positive covariation between handling time and attack rate. Low handling times (equivalent to large maximum consumption rates) and low attack rates create strong regulation (line A in top left panel of Figure 3.2). Large handling times and attack rates result in weaker regulation (line C in top left panel of Figure 3.2). Strong regulation causes the population to return to the fixed point rapidly, and reduces the frequency of large fluctuations caused by demographic effects of predation. These demographic effects are independent of environmental stochasticity and become relevant for extinction risk only for small populations sizes (Lande 1993; Lande, Steinar & Sæther 2004).

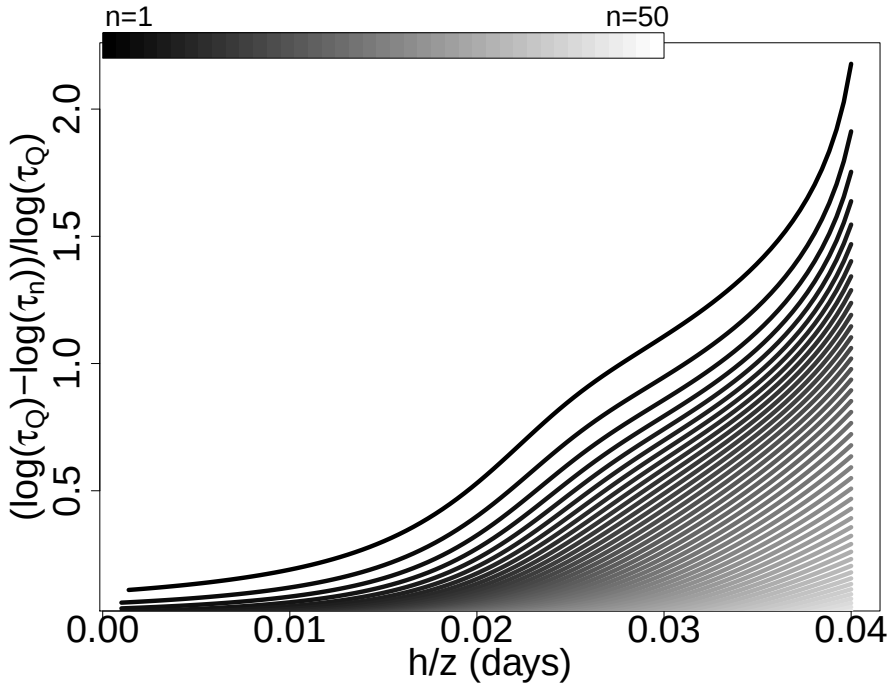

Figure 3.7: Relative difference between the logarithm of the MTE obtained when the QSD is set as initial distribution (τ_Q) and the logarithm of the MTE when $\delta(n)$ is chosen as initial distribution (τ_n), as a function of handling time over predator abundance, for type II functional response when the equilibrium density is kept fixed at $n_0 = 50$.

In (Nåsell 2001) and (Assaf & Meerson 2010) an approximate expression of the quasistationary distribution and the mean time to extinction is derived for the stochastic logistic model and the SIS model of epidemics, both of these are slight simplifications of our model. Our result could be investigated more analytically using refined approximation techniques (Ovaskainen & Meerson 2010). The bistability emerging for high values of handling time in our study, for type III functional response, could also be investigated using approximation techniques.

Our results and main conclusion are robust to both random and periodic fluctuations of predator abundance through time. That is, the qualitative result

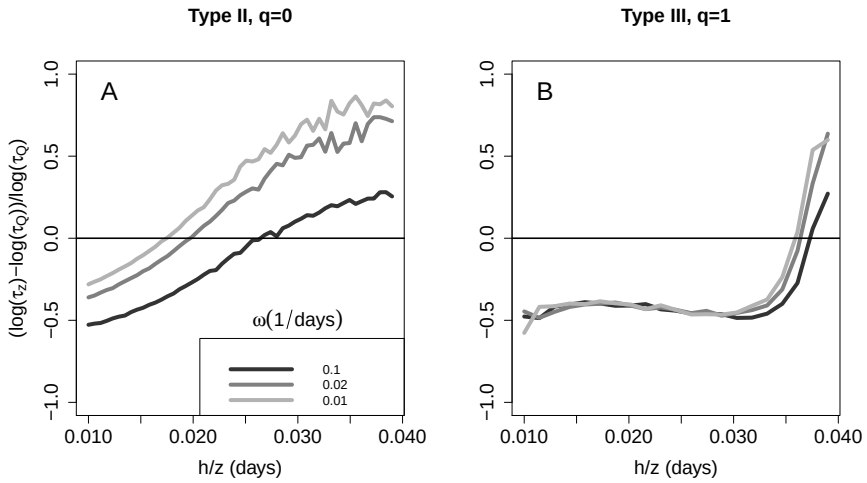

Figure 3.8: Relative difference between the logarithm of the MTE obtained numerically with a predator seasonality given by expression (3.13) (τ_z) and the logarithm of the MTE when the predator abundance is fixed (τ_Q), as a function of handling time over predator abundance. Simulations are done for type II (A) and type III (B) functional response with the foraging parameters of the model with fixed predator abundance when the equilibrium prey population size is kept fixed at $n_0 = 50$ and the QSD is set as initial condition (100 replicates).

remains; changes in the foraging parameters of the predator affect the MTE of the prey, with increases in handling time causing a decrease in MTE. Note that these results were obtained by numerical simulations, therefore, they are limited to the range of parameters explored and are, thus, less general than the other analytical results presented in the paper. Understanding more generally, and in more detail, how and why variability in predator and prey growth rate affect the MTE would be a natural avenue for future research.

Another opportunity for future research is to investigate the consequences of feedbacks between prey and predator abundance by, for example, investigating a bivariate stochastic predator-prey model model. Such a model could allow investigation of how the MTE depends on the strength of this feedback, which itself could result from the extent of specialism / generalism of the predator (specialist predators are likely to have stronger feedbacks Turchin (2003)). Extinction dynamics has been studied for a stochastic predator-prey model with linear interaction rates (Parker & Kamenev 2009). It would be interesting to extend this work to include non-linear functional responses.

Further investigations could constrain analyzes to regions of parameter space and combinations of foraging parameters that occur in reality. Introducing allometric relationships between foraging parameters and relating them to the growth parameters would be one way to constrain such analysis. We have shown this dependence as a function of the stable equilibrium density in equation (3.10). This relation can be generalized using allometric scaling relations between attack rate and handling time (Brose, Williams & Martinez 2006). Such an allometric scaling would require a more general formulation of the model, including predator biomass and prey biomass as other parameters.

This work has application both to studies on extinction risk and to studies on foraging theory. Most of the existing theoretical studies about complex communities do not incorporate the effects of demographic stochasticity and use deterministic measures of persistence to assess the extinction risk (Brose *et al.* 2004; Hofbauer, Kon & Saito 2008; Dunne & Williams 2009; Sahasrabudhe

& Motter 2011). We focused on the strong dependence of extinction times on foraging parameters in order to stress the relevance of interspecific interactions. Our approach may lead to new insights into the determinants of extinctions and can be used to increase the predictive understanding of extinction processes.

3.6 Supplementary information

3.6.1 Analysis of deterministic equation

We derive an adimensional formulation of (3.6) to simplify our subsequent analyzes. We scale the number of individuals with the maximum population size $m = n/k$, and the characteristic time with the intrinsic birth rate $\tau = \lambda t$. With these rescalings, the adimensional form of equation (3.6) is

$$\frac{dm(\tau)}{d\tau} = m(I - m) - \frac{m^{q+1}}{a + bm^{q+1}}. \tag{3.14}$$

We next study the fixed points of (3.6) and analyze their local stability. Putting $dm/d\tau = 0$ in equation (3.14), the fixed points of the system are the extinction state $m = 0$ and the solutions of

$$bm^{q+2} - bIm^{q+1} + m^q + am - Ia = 0, \tag{3.15}$$

where $I = 1 - 1/R_0$, and $R_0 = \lambda/\mu$ is the basic reproductive ratio, and where we introduced the following adimensional parameters:

$$a = \frac{\lambda}{\alpha k^q z}, \quad b = \lambda h \frac{k}{z}. \tag{3.16}$$

From the right lower panel of figure 3.2, we see that equation (3.15) has at most three real and positive solutions. Note that there will be positive solutions only when $0 < R_0 < 1$ i.e. when $\lambda < \mu$.

When $q = 0$ (type II functional response) equation (3.15) simplies to

$$bm^2 + (a - bI)m + (1 - aI) = 0, \tag{3.17}$$

which has real solutions when $(a + bI)^2 > 4b$. If this condition is satisfied then the two real solutions of (3.17) are

$$m_{1,2} = \left(\frac{(bI - a) \pm \sqrt{(a + bI)^2 - 4b}}{2(1 - aI)} \right). \tag{3.18}$$

We now proceed to linear stability analysis of the fixed points of equation (3.14). The Jacobian of equation (3.14) is

$$J(m) = \frac{d}{dm}\left(\frac{dm}{dt}\right) = I - 2m - \frac{(q + 1)am^q}{(a + bm^{q+1})^2}, \tag{3.19}$$

and putting $m = 0$ in (3.19) implies that

- if $q = 0$ then $J(0) = r - \alpha z$ where $r = \lambda - \mu$ so
  1. if $r < \alpha z$ the extinction state is stable
  2. if $r > \alpha z$ the extinction state is unstable
- and if $q > 0$ the extinction state is always unstable

This stability analysis of the extinction state leads to the two extinction scenarios described in the Methods section.

3.6.2 Mean time to Extinction

Here we present the mathematical and numerical tools needed to obtain the quasi-stationary distribution and the mean time to extinction of a general birth and death process when ultimate extinction is certain (given sufficient time). In (Nisbet & Gurney 1982) there is a broader presentation of the methods presented hereafter. We define the process as the time evolution of a random variable $\{X(t), t \geq 0\}$ in a finite state space $\{0, 1, \cdots, k\}$ where

the origin is an absorbing barrier (Nisbet & Gurney 1982). A convenient notation for the master equation (3.4) is

$$\frac{d\mathbf{p}(t)}{dt} = \mathbf{p}(t)\mathbf{A}, \tag{3.20}$$

where $\mathbf{p}(t) = (p(0,t), p(1,t), \cdots, p(k,t))$ is the row vector containing the state probabilities and the matrix $\mathbf{A}$ contains the transition rates as follows:

$$\mathbf{A} = \begin{pmatrix} -G(0) & B(0) & 0 & \cdots & 0 \\ D(1) & -G(1) & B(1) & \cdots & 0 \\ 0 & D(2) & -G(2) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & -G(k) \end{pmatrix}, \tag{3.21}$$

with $G(n) = B(n) + D(n)$. $\mathbf{A}$ is a tridiagonal matrix in which all row sums equal 0. Note also that the first row is a row of zeros when using the rates (3.3). The solution of the master equation (3.4) will give the probability of having n individuals at time t; in other words $p(n,t) = Prob\{X(t) = n\}$.

We now define $T(n)$ and $R(n)$ as

$$\begin{aligned} T(n) &= \frac{B(1)B(2)\cdots B(n-1)}{D(2)D(3)\cdots D(n)}, \\ R(n) &= \frac{B(1)B(2)\cdots B(n-1)}{D(1)D(2)\cdots D(n-1)}. \end{aligned} \tag{3.22}$$

Note that $T(n) = R(n)\frac{D(1)}{D(n)}$. Next we partition the state space of the original process into two subsets, $\{0\}$ and $Q = \{1, 2, \cdots, k\}$. Q is the set of transients for $\{X(t)\}$ while $\{0\}$ is the absorbing state for $\{X(t)\}$. Correspondingly we can partition the state vector $\mathbf{p}(t)$ and the transition matrix $\mathbf{A}$, and obtain from equation (3.20)

$$\left[\frac{dp(0,t)}{dt}; \frac{d\mathbf{p}_Q(t)}{dt}\right] = [p(0,t); \mathbf{p}_Q(t)] \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{a} & \mathbf{A_Q} \end{pmatrix}. \tag{3.23}$$

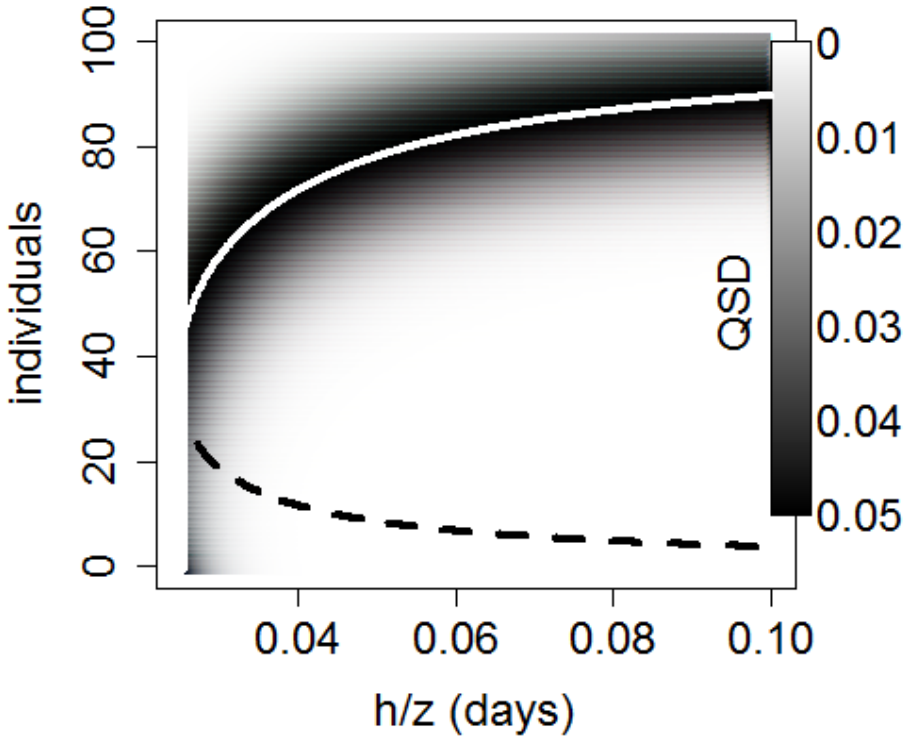

Figure 3.9: Quasistationary distribution (QSD) of the model (3.4) as a function of handling time over predator abundance for type II functional response at fixed attack rate per predator abundance $\alpha z = 1.5d^{-1}$. The continuous line represents the stable fixed point while the dashed line represents the unstable fixed point of the model. The other parameters are as specified in the legend of figure 3.1

Here $\mathbf{p}_Q(t)$ is the vector of probabilities in the transient states and $\mathbf{a} = (D(1), 0, \cdots, 0)^T$. With this separation we can split the master equation (3.20) into:

$$\begin{aligned} dp(0,t)/dt &= \mathbf{p}_Q(t)\mathbf{a} = D(1)p(1,t), \\ d\mathbf{p}_Q(t)/dt &= \mathbf{p}_Q(t)\mathbf{A_Q}. \end{aligned} \tag{3.24}$$

Before absorption (extinction) the process takes values in the set of the transients. As in equation (3.7) we define the conditional probability $p_c(n,t) = P\{X(t) = n | X(t) > 0\}$ of having n individuals at time t knowing absorption (extinction) has not occurred and, using equations (3.24), this can be expressed as

$$\mathbf{p}_c(t) = \frac{\mathbf{p}(t)}{1 - p(0,t)} = \frac{\mathbf{p}_Q(t)}{1 - p(0,t)}. \tag{3.25}$$

Differentiating equation (3.25) and using the master equation (3.4) and equations (3.24) we obtain an equation for $\mathbf{p}_c(t)$:

$$\begin{aligned} \frac{d\mathbf{p}_c(t)}{dt} &= \frac{d\mathbf{p}_Q}{dt}\left(\frac{1}{1-p(0,t)}\right) + \frac{\mathbf{p}_Q(t)}{(1-p(0,t))^2}\frac{dp(0,t)}{dt} \\ &= \mathbf{p}_c(t)\mathbf{A_Q} + D(1)p_c(1,t)\mathbf{p}_c(t). \end{aligned} \tag{3.26}$$

Setting the the right-hand side of expression (3.26) equal to zero we obtain an equation for the quasi-stationary distribution $\boldsymbol{\pi} = (\pi(1), \pi(2), \cdots, \pi(k))$, defined as the distribution of the transient states conditioned on the fact that there has not yet been extinction:

$$\boldsymbol{\pi}\mathbf{A_Q} = -D(1)\pi(1)\boldsymbol{\pi}. \tag{3.27}$$

In other words the quasistationary distribution $\boldsymbol{\pi}$ is the left eigenvector of $\mathbf{A}_Q$ with eigenvalue $-D(1)\pi(1)$.

It can be shown that $\pi(n)$ satisfies the recursive formula:

$$\pi(n) = T(n)\sum_{i=1}^{n} \frac{(1 - \sum_{j=1}^{i-1}\pi(j))}{R(i)}\pi(1). \tag{3.28}$$

Once $\pi(1)$ is known then $\pi(2), \pi(3), \ldots, \pi(n)$ can be determined iteratively. But $\pi(1)$ can only be obtained by knowing all other elements to $\sum_n \pi(n) = 1$. For this reason, the analytic determination of $\boldsymbol{\pi}$ is limited to birth death processes with linear transition rates and this is not our case. However, there is an iterative method that can be used to derive numerical approximations for the quasistationary distribution of our process:

- Start with an initial guess for $\pi(1)$.
- Obtain all the $\pi(n)$ using the (3.28) and compute $S = \sum_n \pi(n)$.
- Start the iteration again with $\pi^I(1) = \pi(1)/S$ and obtain $\pi^I(n)$.
- Repeat the process until $\|\pi^{K+1}(n) - \pi^K(n)\| < \delta$. The value δ gives the precision of the algorithm.

Figure 3.9 shows example results from implementing this procedure to derive the QSD of the birth-death process (3.4) with a type II functional response.

The time to extinction τ[3] is a random variable that depends on the initial distribution of the process (Nåsell 2001). We call τ_Q the time to extinction of the birth-death process when the quasi stationary distribution $\boldsymbol{\pi}$ is set as an initial condition, and τ_n is the time to extinction when the initial condition is $X(0) = n$ i.e., when $p(n, 0) = 1$. If absorption has occurred at time t then the events $\{\tau < t\}$ and $\{X(t) = 0\}$ are identical:

$$P\{\tau < t\} = P\{X(t) = 0\} = p(0, t). \tag{3.29}$$

Once the quasistationary distribution is known then the MTE[4] is given by expression (3.9).

The explicit expression of the time to extinction with an arbitrary initial condition is more difficult to obtain. It is a standard result for birth death

[3]this τ is different from the adimensional time used in appendix A

[4]note that with our notation $MTE = E(\tau_Q)$.

processes theory (Nisbet & Gurney 1982) that the expectation of τ_n can be determined explicitly when $X(0) = n$:

$$E(\tau_n) = \frac{1}{D(1)} \sum_{i=1}^{n} \frac{1}{R(i)} \sum_{j=i}^{k} T(j). \tag{3.30}$$

Moreover, the expected time to extinction for an arbitrary initial distribution $\{p(n, 0)\}$ can be derived from (3.30):

$$E(\tau) = \frac{1}{D(1)} \sum_{j=1}^{k} T(j) \sum_{i=1}^{j} \frac{1}{R(i)} \sum_{n=i}^{k} p(n, 0), \tag{3.31}$$

with the assumption that the initial distribution is supported on the set of the transient states.

Bibliography

Abrams, P. A., & Ginzburg, L. R. (2000) The nature of predation: prey dependent, ratio dependent or neither? *Trends in Ecology & Evolution*, **15**, 337-341.

Allesina, S. & Tang, S. (2012) Stability criteria for complex ecosystems. *Nature*.

Assaf, M. & Meerson, B. (2010) Extinction of metastable stochastic populations. *Physical Review E*, **81**, 1-18.

Bartlett, M. S. (1960) *Stochastic Population Models in Ecology and Epidemiology*. Methuen, London.

Bascompte, J. & Solé, R. V. (1995) Rethinking complexity: modelling spatiotemporal dynamics in ecology. *Trends in Ecology & Evolution*, **10**, 361-366.

Beissinger, S. R. & McCullough, D. R. (2002) *Population viability analysis*. University of Chicago Press, Chicago.

Black, A. J., & McKane A. J. (2012) Stochastic formulation of ecological models and their applications. *Trends in Ecology & Evolution*, 1-9.

Brook, B. W., J. J. O'Grady, A. P. Chapman, M. A. Burgman, H. R. Akçakaya, & Frankham, R. (2000) Predictive accuracy of population viability analysis in conservation biology. *Nature*, **404**, 385-7.

Brose, U., Ostling, A., Harrison, K. & Martinez, N. D. (2004) Unified spatial scaling of species and their trophic interactions. *Nature*, **428**, 167-171.

Brose, U., Williams, R. J. & Martinez, N. D. (2006) Allometric scaling enhances stability in complex food webs. *Ecology Letters*, **9**, 1228-1236.

Drossel, B., McKane, A. J. & Quince, C. (2004) The impact of nonlinear functional responses on the long-term evolution of food web structure. *Journal of Theoretical Biology*, **229**, 539-548.

Dunne, J. A. & Williams, R. J. (2009) Cascading extinctions and community collapse in model food webs. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **364**, 1711-1723.

Ebenman, B., Law, R. & Borrvall, C. (2004) Community viability analysis: the response of ecological communities to species loss. *Ecology*, **85**, 2591-2600.

Foley, P. (1994) Predicting extinction times from environmental stochasticity and carrying capacity. *Conservation Biology*, **8**, 124-137.

Gardinier, C. (2009) *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Springer, Berlin.

Gillespie, D. T. (1977) Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, **81**, 2340-2361.

Hammill, E., Petchey, O. L. & Anholt, B. R. (2010) Predator functional response changed by induced defenses in prey." *The American Naturalist*, **176**, 723-731.

Hanski, I., Hansson, L. & Henttonen, H. (1991) Specialist predators, generalist predators, and the microtine rodent cycle. *Journal of Animal Ecology*, **60**, 353-367.

Hofbauer, J., Kon, R. & Saito, Y. (2008) Qualitative permanence of Lotka-Volterra equations. *Journal of Mathematical Biology*, **57**, 863-881.

Holling, C. S. (1959) Some characteristic of simple types of predation and parasitism. *The Canadian Entomologist*, **XCI**, 385-398.

Holling, C. S. (1965) The functional response of predators to prey density. *Memoirs of the Entomological Society of Canada*, **45**, 1-60.

Jeschke, J. M., Kopp, M. & Tollrian, R. (2002) Predator functional responses: discriminating between handling and digesting prey. *Ecological Monographs*, **72**, 95-112.

Jeschke, J. M., Kopp, M. & Tollrian, R. (2004) Consumer-food systems: why type I functional responses are exclusive to filter feeders. *Biological Reviews*, **79**, 337-349.

Kramer, A. M. & Drake, J. M. (2010) Experimental demonstration of population extinction due to a predator-driven Allee effect. *The Journal of Animal Ecology*, **79**, 633-639.

Lande, R. (1993) Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *The American Naturalist* **142**, 911-927.

Lande, R., Steinar, E. & Sæther, B. E. (2004) *Stochastic population dynamics in ecology and conservation*, Oxfrord series in Ecology and Evolution.

Liebhold, A., & Bascompte, J. (2003) The Allee effect, stochastic dynamics and the eradication of alien species. *Ecology Letters*, **84**, 468-140.

Mace, G. M., Collar, N. J., Gaston, K. J., Hilton-Taylor, C., Akçakaya, H. R., Leader-Williams, N., Milner-Gulland, E. J. & Stuart, S. N. (2008) Quantification of extinction risk: IUCN‘s system for classifying threatened species. *Conservation Biology*, **22**, 1424-1442.

May, R. M. (2001) *Stability and complexity in model ecosystems*. Princeton University Press, Princeton, New Jersey.

McCann, K. S. (2000) The diversity-stability debate. *Nature*, **405**, 228-233.

McKane, A. J., & Newman, T. J. (2004) Stochastic models in population biology and their deterministic analogs. *Physical Review E*, **70**, 1-19.

McKane, A. J., & Newman, T. J. (2005) Predator-prey cycles from resonant amplification of demographic stochasticity. *Physical Review Letters*, **94**, 1-4.

Nåsell, I. (2001) Extinction and quasi-stationarity in the Verhulst logistic model. *Journal of Theoretical Biology*, **211**, 11-27.

Nisbet, R. M. & Gurney, W. C. S. (1982) *Modelling fluctuating population.* The Blackburn Press.

Ovaskainen, O. & Meerson, B. (2010) Stochastic models of population extinction. *Trends in Ecology & Evolution*, **25**, 643-652.

Parker, M. & Kamenev, A. (2005) Extinction in the Lotka-Volterra model. *Physical Review E*, **80**.

Powell, C. R. & Boland, R. (2009) The effects of stochastic population dynamics on food web structure. *Journal of Theoretical Biology*, **257**, 170-180.

Prowse, T. A. A., Johnson, C. N., Lacy, R. C., Bradshaw, C. J. A., Pollak, J. P., Watts, M. J. & Brook, B. W. (2013) No need for disease: testing extinction hypotheses for the thylacine using multi-species metamodels. *Journal of Animal Ecology*, **82**, 355-364.

Purvis, A., Gittleman, J. L., Cowlishaw, G. & Mace, G. M. (2000) Predicting extinction risk in declining species. (2000) *Proceedings of the Royal Society B: Biological Sciences*, **267**, 1947-1952.

Rall, B. C., Brose, U., Hartvig, M., Kalinkat, G., SchwarzmÃ$\frac{1}{4}$ller, F., Vucic-Pestic, O. & Petchey, O. L. (2012) Universal temperature and body mass scaling of feeding rates. *Philosophical Transactions of the Royal Society*, **367**, 2923-2934.

Real, L. A. (1977) The kinetics of functional response. *The American Naturalist*, **111**, 289-300.

Sabo, J. L. & Gerber, L. R. (2007) Predicting extinction risk in spite of predator-prey oscillations." Ecological Applications **17**, 1543-1554.

Sabo, J. L. (2008) Population viability and species interactions: Life outside the single-species vacuum. *Biological Conservation*, **141**, 276-286.

Sahasrabudhe, S. & Motter, A. E. (2011) Rescuing ecosystems from extinction cascades through compensatory perturbations. *Nature Communications*, **170**, 1-21.

Smout, S., Asseburg, C., Matthiopoulos, J., Fernández, C., Redpath, S., Thirgood, S. & Harwood, J. (2010) The functional response of a generalist predator. *PloS One*, **5**, e10761.

Spitze, K. (1985) Functional response of an ambush predator: *Chaoborus americanus* predation on *Daphnia Pulex*. *Ecology*, **66**, 938-949.

Turchin, P. (2003) *Complex population dynamics.* Princeton University Press, Princeton, New Jersey.

Vucic-Pestic, O., Birkhofer, K., Rall, B. C., Scheu, S. & Brose, U. (2010) Habitat structure and prey aggregation determine the functional response in a soil predator-prey interaction. *Pedobiologia*, **53**, 307-312.

Williams, R. J. & Martinez, N. D. (2004) Stabilization of chaotic and non-permanent food-web dynamics. *The European Physical Journal B*, **38**, 297-303.

Witting, L., Tomiuk, J. & Loeschcke, V. (2000) Modelling the optimal conservation of interacting species. *Ecological Modelling*, **125** 123-144.

Chapter 4

Extinction rates in an experimental microbial aquatic community

Gian Marco Palamara, Nicholas Worsfold & Owen L. Petchey

4.1 Abstract

Extinction is a natural phenomenon and has occurred throughout the history of life. The current rate of extinction of species is becoming comparable to those observed during previous mass extinction events. Therefore, it is critically important to improve our predictive understanding of the causes and consequences of biodiversity loss. Most of the models used to asses extinction risk are based on single species population dynamics and do not include explicitly the effects species interactions on the extinction process.

We present data from a highly replicated microcosm experiment where the extinction times of 17 freshwater protists species forming a small food web was recorded. Four different treatment combinations have been used, varying temperature and nutrient concentration in the microcosms. We use survival analysis to show how interspecific dependencies and environmental factors (energy and temperature), influence the trajectory of ecosystem failure that results from species extinctions.

Environmental factors are the most significant in predicting the extinction times distribution of the species in the food web. We found no significant effect of body size on species extinction times both among the same trophic groups and across the whole community. We also found that the number of competitors can be used to improve the prediction of the extinction time distributions.

The experiment shows a clear interplay between species richness, environmental factors, and species interactions. The use of simple phenomenological models is a first step in disentangling the effect of both the environment and species interactions on the timing of extinction events. As climate change continues to escalate, it is becoming critically important to assess those effects. In the future, we aim to generalize our approach by relating survival analysis to more mechanistic models of population dynamics.

Key Words: Microcosm Experiments, Protists, Survival Analysis, Species Interactions, Food Webs

4.2 Introduction

Ecosystems are losing species at rates that have lead to prophecies of a sixth mass extinction event (Pimm *et al.* 2014). Estimates vary but include current extinction rates up to one thousand times faster than in the fossil record. Future rates could be ten times faster again (MEA 2005). Species are the building blocks of ecosystems and exist in communities dynamically interrelated with non-living components. The living and non-living components affect each other in complex exchanges of energy and matter, performing functions such as nutrient recycling, photosynthesis, and water purification (Chapin *et al.* 2000). As a consequence, the loss of species is compromising the ability of ecosystems to perform their functions (Cardinale *et al.* 2012) and put at risk the services on which humanity relies (Daily 1997; Balvanera *et al.* 2006).

Research about the functional consequences of extinctions has revealed two key findings (Balvanera *et al.* 2006; Cardinale *et al.* 2006). First, a larger number of extinctions cause, on average, a larger reduction in ecosystem functioning. Second, variation around this average can be substantial and is caused by the importance of species identity. The influence of species identity is further reflected in research about how the order in which species are lost influences the consequences of this loss (Petchey & Gaston 2002; Larsen, Williams & Kremen 2005; Gross & Cardinale 2005; Zavaleta & Hulvey 2004; Srinivasan *et al.* 2007). For example, loss of larger species first, followed by smaller species, can cause more rapid loss of ecosystem functioning in marine sediments than do random extinction orders, according to simulation based predictions (Solan *et al.* 2004). This means that knowing how many species might suffer extinction is insufficient; we must also know when particular species will go extinct.

Research about the risk and probable timing of extinction largely considers species as isolated entities (Sabo 2008) influenced by an abiotic environment (Belovsky 1999; Lande 1993; Drake & Lodge 2004; Purvis *et al.* 2000; Petchey *et al.* 1997; Mace & Lande 1991) and has demonstrated the role of species traits, stochasticity, genetic effects, environmental change, and migration in species' extinction probabilities (e.g. Lande (1993); Hanski (1998); Cardillo *et al.* (2005); Pimm *et al.* (1988); Gaston & Blackburn (1995)). This species-centric approach has yielded many important insights including trait-based extinction scenarios in models of realistic simulations of biodiversity loss (Petchey & Gaston 2002; Larsen, Williams & Kremen 2005; Gross & Cardinale 2005; Zavaleta & Hulvey 2004; Srinivasan *et al.* 2007; Solan *et al.* 2004).

Species exist in communities, however, linked by often complex networks of interactions, such as competition, predation, and mutualisms. One consequence of these interactions is dependencies between species' extinctions, a phenomenon clearly reflected when the loss or removal of one species causes the loss of other species (e.g., Paine (1966)). Theoretical models also show that a first extinction can cause a cascade of secondary extinctions, again, clear evidence of the presence of dependencies between species' extinction risks (e.g., Dunne & Williams (2009)). Therefore, there is clear evidence that interspecific dependencies can affect the occurrence, rate, and order of species' extinctions (Koh *et al.* 2004).

Models to assess extinction risks in communities depend on the spatial (Rybicki & Hanski 2013) and temporal scale (Raup & Sepkoski 1984; Newman 1996) of the system. Studies including "realistic" extinction orders have been limited to trait-based, species-centric, assumptions about extinction orders (Petchey & Gaston 2002; Larsen, Williams & Kremen 2005; Gross & Cardinale 2005; Zavaleta & Hulvey 2004; Srinivasan *et al.* 2007). Recently models of extinction risk have been developed in the framework of community ecology and have provided insights on the mechanisms behind community collapse (Ebenman, Law & Borrvall 2004; Ebenamn & Jonsson 2005; Ingram

& Steel 2010). These models shed light on the relation between species richness and secondary extinction events, showing how stochastic effects (e.g. demographic stochasticity) play a fundamental role in shaping the structure of ecological communities after a first extinction event takes place (Ebenman, Law & Borrvall 2004). More mechanistic models (e.g. (Ebenman, Law & Borrvall 2004) or Powell & Roland (2009)) take directly into account the determinants of population dynamics such as growth rates and feeding rates together with the demographic and environmental factors affecting population dynamics, while more phenomenological models (e.g. Ingram & Steel (2010)) focus on characterizing the global properties of the extinctions in relation to environmental and demographic factors.

Here we propose a new phenomenological framework to further test the effect of environmental factors and interspecific interactions on the extinction rates of species in a community. We aim to provide a community-centric perspective in the study of extinction risk using the statistical tools of survival analysis (Hougaard 2000; Kleinbaum & Klein 2005). We use survival analysis on data coming from a highly replicated experimental study of extinction dynamics in a complex laboratory community (17 species at multiple trophic levels with four environmental treatments) (Worsfold 2007). The data from the experiment show evidence of the dependency of the extinction rates on the extinction (or not) of the other species. These dependencies are driven by environmental factors and interspecific interactions. The main aim of this study is to gain insight about the combined effect of interspecific interactions and environmental factors on the determinants of extinction events. We found that including environmental factors (i.e. Temperature and nutrient concentration) in survival analysis improves the prediction of extinction times of the species in the food web. While including species' interactions (i.e. predation and competition) provides less improvement to the predictions.

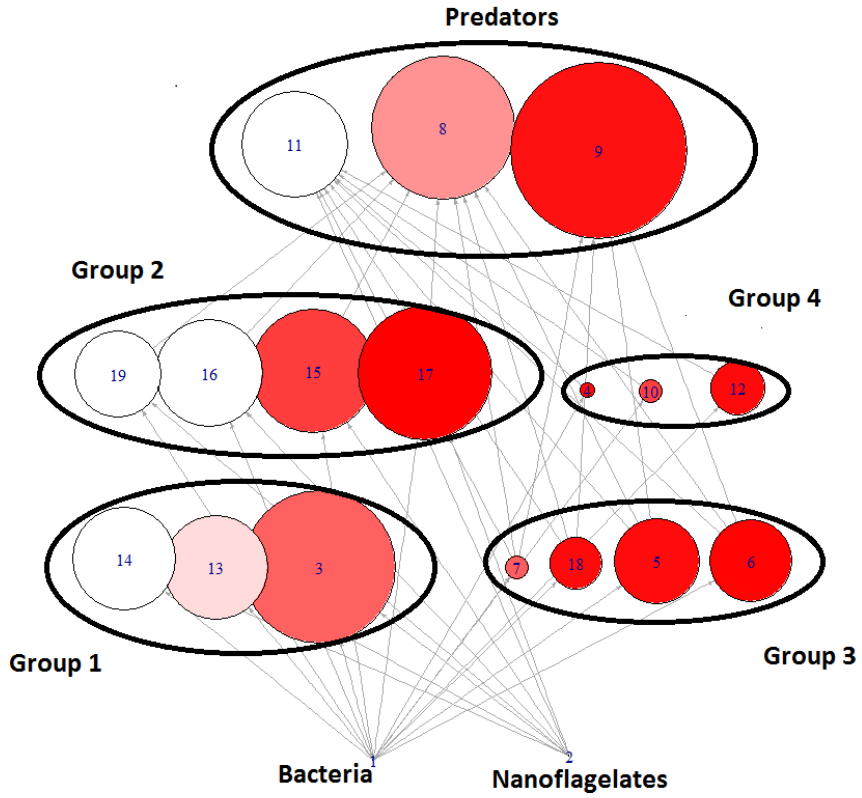

Figure 4.1: Food web: species are grouped according to their trophic group. Each node has a size proportional to the logarithm of the body size of the species and a color indicating the mean extinction time of the species ($< t >$) computed by maximizing likelihood 4.1 across all treatments. The arrows correspond to a trophic link from the prey to the predator. Species' mean extinction time are listed below: 3 *Blepharisma japonicum* ($< t >= 7.5 \pm 2$ weeks), 4 *Chilomonas paramecium* ($< t >= 0.6 \pm 0.5$ weeks), 5 *Colpidium striatum* ($< t >= 0.8 \pm 0.5$ weeks), 6 *Colpoda cucculus* ($< t >= 0.3 \pm 1$ weeks), 7 *Cyclidium glaucoma* ($< t >= 7.35 \pm 4$ weeks), 8 *Didinium nasutum* ($< t >= 11.4 \pm 3.5$ weeks), 9 *Dileptus anser* ($< t >= 1.4 \pm 0.8$ weeks), 10 *Entosiphon* ($< t >= 5 \pm 2.3$ weeks), 11 *Euplotes patella* ($< t >\sim \infty$), 12 *Loxocephallus* ($< t >= 0.8 \pm 0.5$ weeks), 13 *Lepadella* ($< t >= 17.2 \pm 3.7$ weeks), 14 *Dicanophoridae* ($< t >\sim \infty$), 15 *Paramecium bursaria* ($< t >= 4.7 \pm 3.2$ weeks), 16 *Paramecium caudatum* ($< t >\sim \infty$), 17 *Telotrochiudium* ($< t >= 0$ weeks), 18 *Tetrahymena piriformis* ($< t >= 0.8 \pm 0.8$ weeks), 19 *Vorticella* ($< t >= 23 \pm 5.5$ weeks).

4.3 Methods

4.3.1 Experimental setting

The experiment was conducted by Nicholas Worsfold at the University of Sheffield (Worsfold 2007). A factorial design with two levels of chemical energy and two temperatures was used. Microcosms consisted of a clear polystyrene Petri dish (diameter 90mm, height 25mm) containing 50mL of liquid growth medium, 17 taxa of aquatic eukaryotic microorganisms, unknown heterotrophic nanoflagellates and an unknown bacterial flora. Energy was manipulated by adding either 0.275g ("low energy") or 0.825g ("high energy") of ground Protozoan Pellet (a source of organic nutrients; Carolina Biological Supply, Burlington, NC) to one litre of Chalkley's medium (Thompson, Rhodes & Pettman 1988), an inorganic salt solution that provided no energy. The medium was autoclaved before use. Nutrient enrichment increases the production of bacteria, the basal trophic level in this community (Kaunzinger & Morin 2008). Temperature was manipulated using two incubators, one at $15°C$ and one at $20°C$. To remove possible incubator effects, the temperature of each incubator was switched weekly and the microcosms moved so that the two temperature treatments spent equal time in both incubators. There were 50 replicates of each of the four treatments, (LL i.e., $15°C$ low energy, LH i.e., $15°C$ high energy, HL i.e., $20°C$ low energy and HH i.e., $20°C$ high energy), giving 200 microcosms in total. The short generation times of the organisms in the communities (approximately 4-24 hours) allowed us to collect long-term data on the effect of the treatments. All microcosms were kept in the dark at their experimental temperature from the beginning of the community establishment sequence to the end of the nine week experimental period.

Microcosm communities were assembled over a six day period, leading up to the start of the monitoring period (day 0), as follows. On day -6 three species of bacteria (*Bacillus cereus*, *Bacillus subtilis* and *Serratia marcescens*) were added to the two concentrations of sterile media. On day -4 the media were

poured into the experimental Petri dishes and approximately 20 individuals of each of 13 taxa of bactivorous protozoa and rotifers were introduced. Bacterivores were introduced by adding a few drops (between 0.05mL and 0.3mL) of stock cultures to each microcosm, depending on the population density of each species. Unknown bacteria and nanoflagellates present in some stock cultures were also unavoidably added. Nine of the bactivorous species were ciliates; *Cyclidium glaucoma* (Muller), *Telotrochidium henneguyi* (Faure-Fremiet), *Vorticella sp.* (L), *Paramecium caudatum* (Ehrenberg), *Paramecium bursaria* (Ehrenberg), *Colpidium striatum* (Stokes), *Tetrahymena pyriformis* (Ehrenberg), *Colpoda cucculus* (Muller), *Loxocephalus sp.* (Eberhard), two were flagellates; *Entosiphon sp.* (Dujardin), *Chilomonas paramecium* (Ehrenberg) and two were rotifers; *Lepadella sp.* (Bory de St Vincent) and an unknown member of the family Dicanophoridae (henceforth *Dicanophoridae sp.*). On day -2 each of the bacterivorous species were added again and 20 individuals of three predatory ciliates, *Euplotes patella* (Muller), *Didinium nasutum* (Muller) and *Dileptus anser* (Muller) and the omnivorous ciliate *Blepharisma japonicum* (Suzuki) were also introduced into each microcosm. Twenty individuals of the predators and omnivore alone were added for the second time on day 0. This was the "first day" of the experiment and all future times relate to this day. Bacterivores were maintained separately in stock cultures before the experiment at a nutrient concentration of 0.55g of Protozoan Pellet per liter at $20°C$. Of the predators, *Didinium nasutum* was maintained on a diet of *Paramecium caudatum*, *Euplotes patella* on *Chilomonas paramecium* and *Dileptus anser* on *Colpidium striatum*. The ciliate *Blepharisma japonicum*, which is invulnerable to the other predators, can assume two morphs: a bacterivorous morph (which was introduced from stock cultures) and an omnivorous morph (which may develop from the bactivorous morph under certain conditions and can consume most species in our communities, including smaller individuals of its own species). It was therefore added at the same time as the predators, but subsequently, no omnivorous morphs were seen and it may be considered a bacterivore in this experiment. Adding the predators, individual-by-individual, using a micropipette minimized the number of prey items that were introduced with

the predators. The experiment was set up over five days, with ten replicates of each treatment established each day, allowing all microcosms to be sampled after the same amount of time. The community food web is shown in figure 4.1.

Microcosms were sampled every 7 days for 8 weeks by scanning the whole microcosm under a dissecting microscope (magnification range 8-80x) for 5 minutes and recording the presence or absence of each species. Occasionally, a species was recorded as absent but subsequently seen in the following weeks' sample. Therefore, all microcosms from which a species disappeared between weeks 7 and 8 were checked one week later to confirm that the species was indeed absent from the final (week 8) sample. Some species present in the experimental communities can survive unfavorable periods by encysting. However, no species capable of encysting were seen after they disappeared, so we treated any possible encysting as extinction. After sampling, each microcosm was gently swirled to homogenize the media and 10% of the volume was replaced with fresh sterile medium to renew nutrients and avoid the build up of metabolic waste. Evaporative water loss (approximately 1-4% per week) was replaced with distilled water.

4.3.2 Statistical analysis

Data from the experiment is presence or absence of each species for each of the eight weeks of the experiment, for each of the 200 replicates. We measured the week at which each species went extinct in each replicated microcosm. We were then able to estimate the extinction times distribution of all the species across the four environmental treatments (see figure 4.7). We used classic survival analysis (Kleinbaum & Klein 2005) fitting to the observed extinction time distribution a likelihood function given by

$$L_S(\lambda, k) = \prod_i \int_{t_{i-1}}^{t_i} W_S(\lambda, k; x) dx, \tag{4.1}$$

where $W_S(\lambda, k; x)$ is the Weibull distribution, describing the probability that species S is going extinct at time x (measured in weeks) and depends on two parameters i.e., the scale λ and the shape k. The times t_i are $(0, 1, 2, 3, 4, 5, 6, 7, 8, \infty)$ meaning that if a species is not going extinct after week 8 we assume it will go extinct in a time comprised between 8 and ∞. We fitted likelihood 4.1 to the observed extinction times distributions for all the four treatments using function mle2() in the package bbmle of R Bolker (2013). We then computed the predicted mean extinction time $< t_S >= \lambda_S^* \Gamma(1+1/k_S^*)$ with its standard error, where λ_S^* and k_S^* are the scale and shape for species S. We performed the fitting first considering all the data and then separating the fitting by treatment combinations.

We also investigated the predicted mean extinction time of each species as function of the logarithm of the body size of the species, for each treatment. We use a linear model of the form $< t_S >= A + Blog(m_S)$, where m_S is the body size of species S measured in milligrams and A and B are the slope and the intercept of the linear model, using first all treatments and then treatment by treatment and by trophic group.

We extended the survival analysis to investigate the combined effect of the environment and of species' interactions. To do so we used the likelihood function given by 4.1, with the scale and the shape of the Weibull distribution are given by

$$\begin{aligned} \lambda(t) &= \lambda_0 + \lambda_T T + \lambda_E E + \lambda_I N_I(t) \\ k(t) &= k_0 + k_T T + k_E E + k_I N_I(t) \end{aligned} \tag{4.2}$$

where E and T are binary variables for the environmental treatments and $N_I(t)$ is the number of interactions at the extinction week. N_I was set as the number of competitive interactions at the time of its extinction. The number of competitors of a species can be seen as the number of second neighbors of the node representing that species in the food web graph (Newman 2000). Only for predator species (*Euplotes patella*, *Didinium nasutum*,

Dileptus anser), we set N_I as the number of prey i.e., the number of in-links of the species at the time of its extinction (Newman 2000). The likelihood has now six more parameters where (λ_T, k_T), (λ_E, k_E) and (λ_I, k_I) represent respectively the effect of temperature, of nutrient richness and of species interactions on the extinction times distribution. For simplicity, we do not include parameters related to the statistical interaction between the different factors. Note that setting $\lambda_T = \lambda_E = \lambda_I = 0$ and $k_T = k_E = k_I = 0$ in 4.2 we obtain likelihood 4.1.

We compare, for a selected number of species, the difference between mean predicted and observed extinction time using four different models: the "species only model" (given by likelihood 4.1), the "species + environment model" (i.e. using likelihood 4.1 with 4.2 and $\lambda_I = 0$ and $k_I = 0$), the "species + environment + randomized interactions model" (using likelihood 4.1 with 4.2 with $N_I(t)$ given by a random sample from the number of interactions of each species during the experiment) and the "species + environment + real interactions model" (using likelihood 4.1 with 4.2 with $N_I(t)$ given by the actual number of interactions of the species at the extinction week). Doing so we fit four different Weibull distributions to the empirical extinction times' distributions increasing the number of parameters of the fitted Weibull from two ("species only model") to six ("species + environment model") to eight ("species + environment + randomized interactions model" and "species + environment + real interactions model"). Using the model with randomized number of interactions we reproduce a situation where the number of interactions of each species in the food web is not related to their extinction time. Doing so we want to test if the number of interactions at the extinction time is more relevant in determining the extinction time than the number of interactions before the extinction time.

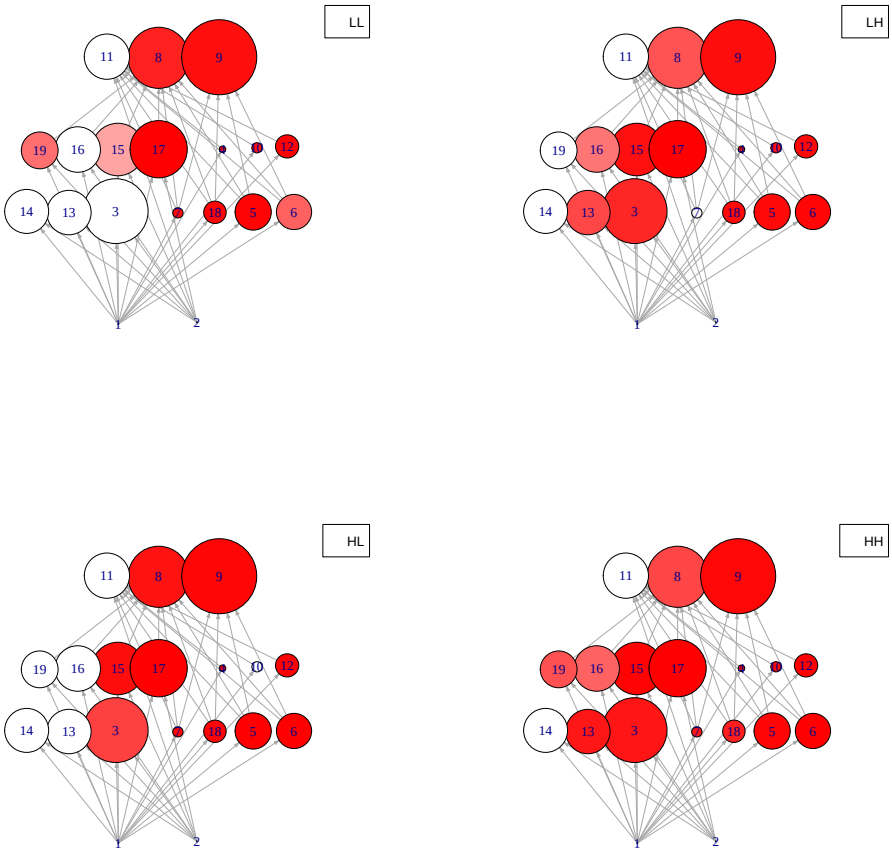

Figure 4.2: food webs with extinction risk (color) and species body size (size of the nodes). Species are grouped according to their trophic group. Each node has a size proportional to the logarithm of the body size of the species and a color indicating the mean extinction time of the species ($< t >$) computed by maximizing likelihood 4.1 treatment by treatments.

4.4 Results

The food web represented in figure 4.1 is composed of 17 species, two trophic levels, bacterivours and predators. At the first trophic level (the bacterivours), we can distinguish four different trophic groups characterized by species having the same predators and the same prey (see figure 4.1). Group 1 is composed by *Blepharisma japonicum*, *Lepadella* and *Dicanophoridae*, they all feed on both bacteria and nanoflagelates and have no predators. Group 2 is composed by *Paramecium bursaria*, *Paramecium caudatum*, *Telotrochidium* and *Vorticella*, they feed on both bacteria and nanoflagelates and are eaten by *Didinium nasutum*. Group 3 is composed by *Colpidium striatum*, *Colpoda cucculus*, *Cyclidium galucoma* and *Tetrahymena piriformis*, they all feed on bateria and are eaten by the three predators *Didinium nasutum*, *Euplotes patella* and *Dileptus anser*. Group 4 is composed by *Chilomonas paramecium*, *Entosiphon* and *Loxocephallus*, they all feed on bacteria and are eaten by *Euplotes patella*. Of the three predators, only *Euplotes patella* feeds on both the bacteriovours and nanoflagelates.

Species richness declined from 17 species to a maximum of eight and a minimum of two species at the end of the experiment (Worsfold 2007). Twelve of the 17 species that were initially introduced were still present in at least one of the 200 microcosms. Only two species *Euplotes patella* and *Dicanophoridae* persisted in all the replicates of the experiment while only one species i.e., *Telotrochidium* always went extinct before the first week of the experiment (Worsfold 2007). Species richness is effected by temperature and energy throughout the experiment and is determined by a significant three-way interaction between temperature, energy and time (see figure 4.3 and for more details see (Worsfold 2007) chapter 5).

The average of the predicted mean extinction times obtained fitting likelihood 4.1 to all the treatments separately is 5.2 weeks for the LL treatment, 4.12 weeks for the LH treatment, 1,35 weeks for the HL treatment and 3.4 weeks for the HH treatment. On average species went extinct faster in treatments

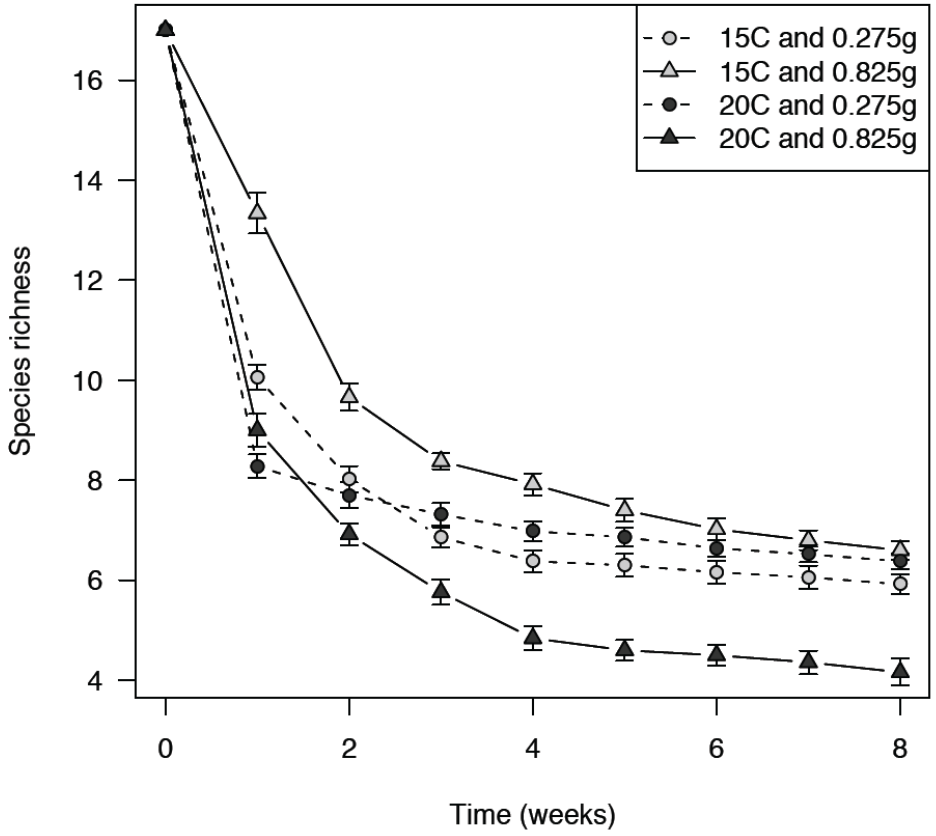

Figure 4.3: mean species richness, treatment by treatment. The error bars show the standard error of species richness in the different treatments.

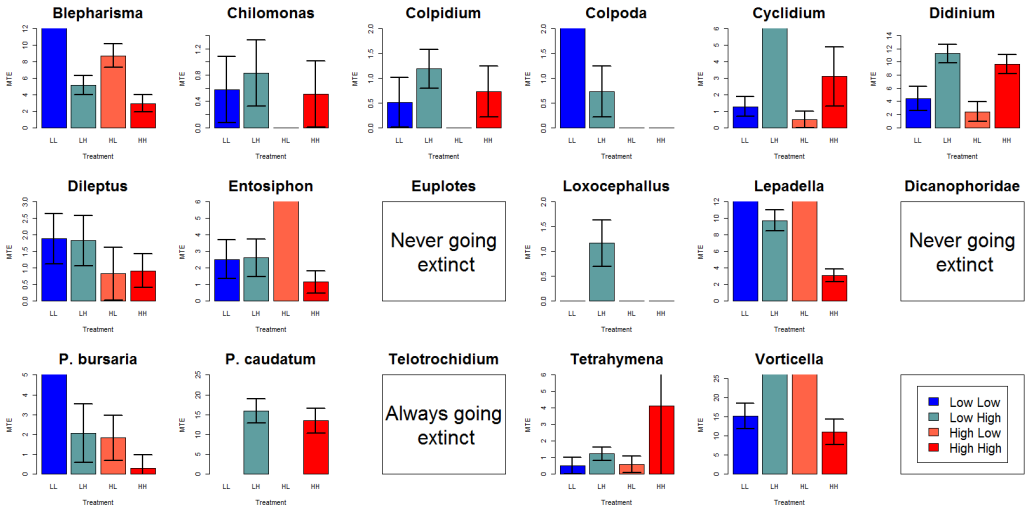

Figure 4.4: Mean time to extinction of the 17 species of the foodweb obtained maximizing likelihood 4.1 treatment by treatment. The four colors refer to the four different treatments as specified in the legend (e.g. Low High refers to the treatment with Low temperature ($15°C$) and high energy ($0.825g$). The error bars show the standard errors of the predicted mean extinction times.

at higher temperature (as shown for example by *Blepharisma*, *P. caudatum* and *Colpoda* in figure 4.4). Species went extinct faster in the high energy treatment (as shown for example by *Entosiphon*, *Lepadella* and *Vorticella* in figure 4.4). In figure 4.2 we see the same effect EXPAND. However, energy has a positive effect on species richness at low temperature (Worsfold 2007). There is, therefore, a combined effect of temperature and enrichment on the extinction times of the species in the food web (for more details see Worsfold (2007), figure 5.1).

Body size played no significant role in determining the extinction determinism. The coefficients of the linear model, fitted when using all treatments are the intercept $A = 4.9 \pm 4.7$ weeks ($p = 0.3$) and the slope $B = 0.37 \pm 1$

weeks/mg ($p = 0.7$). There was no significant correlation also looking at the fit treatment by treatment (see figure 4.5) and across trophic levels.

The difference between predicted an observed extinction week, when using the "species only model" was of maximum five weeks (see figure 4.6 panel A) for the bacterivours and of three weeks for the predators (see figure 4.6 panel B). This difference was reduced by two weeks when including the environmental factors (i.e. using he "species + environment model") for both bacterivours and predators (see figure 4.6). Using the "species + environment + randomized interactions model" did not changed the goodness of the fit, while the difference between predicted an observed extinction week was further reduced up to maximum one week when also competitive interactions are included (i.e. using the "species + environment + ral interactions model" for the bacterivours). We did not observe a relevant effect of the number of interactions on the predicted mean time to extinction except for one species. *Paramecium caudatum* that was the only species that really benefited from including the number of competitors at the extinction weeks as a parameter of the fitted extinction times distribution (see figure 4.6 panel A). On the other hand including the number of prey as interactions, lead to no relevant improvement on the prediction of the extinction time for the predators (see figure 4.6 panel B).

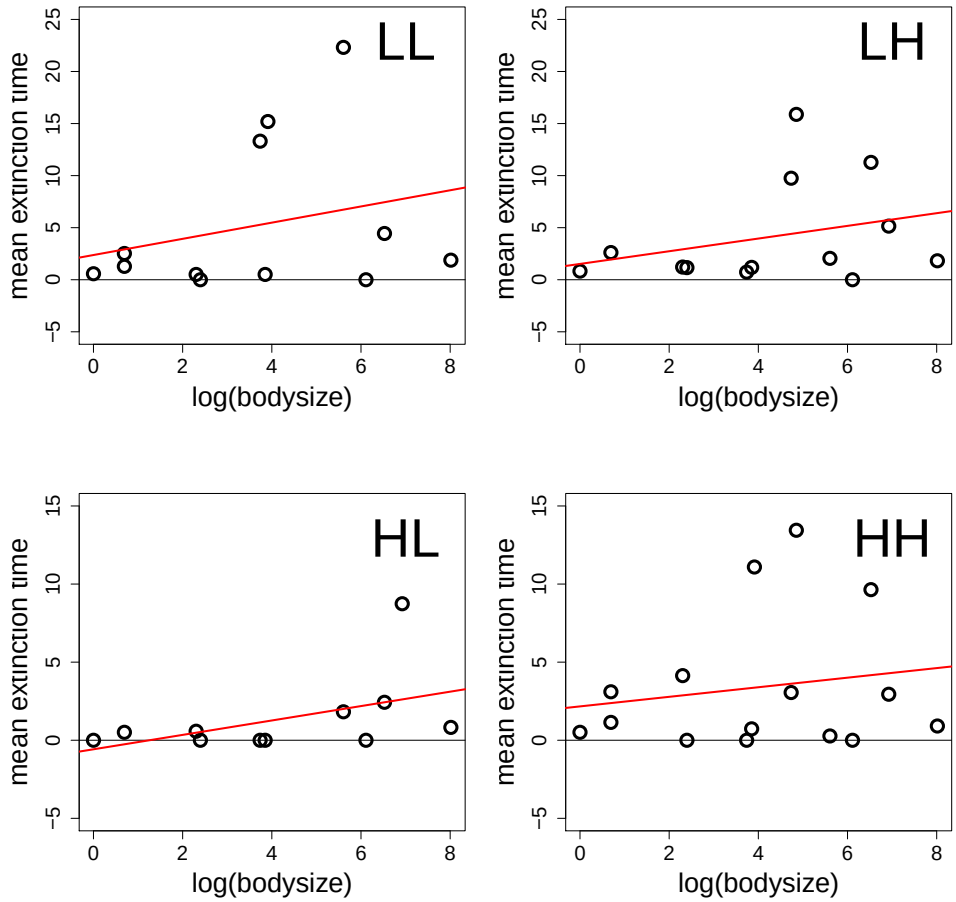

Figure 4.5: Mean time to extinction as a function of the logarithm of the body size o the species for the four different treatment combinations. Respectively top left panel, low temperature low energy (LL), $A = 0.7 \pm 0.9$ weeks ($p = 0.404$), $B = 2.3 \pm 3.9$ weeks/mg ($p = 0.558$); top right panel, low temperature, high energy (LH), $A = 0.6 \pm 0.6$ weeks ($p = 0.32$), $B = 1.5 \pm 2.9$ weeks/mg ($p = 0.6$); bottom left panel, high temperature low energy (HL), $A = -0.57 \pm 1.4$ weeks ($p = 0.7$), $B = 0.46 \pm 0.3$ weeks/mg ($p = 0.142$); bottom right panel high temperature, high energy (HH), $A = 2.16 \pm 2.3$ weeks ($p = 0.36$), $B = 0.3 \pm 0.5$ weeks/mg ($p = 0.54$).

4.5 Discussion

We find that species' interactions are less important than environmental factors (nutrient concentration and temperature) in determining the extinction times distribution of the species in the microbial food web. This finding might be related to the phenomenological model we use to assess the effect of species' interactions on the extinction times distribution. In fact, we don't assess directly the effect of temperature and nutrient concentration on species' interactions, but only their separate effect on species extinction times. Therefore, we exclude potential interactions between population dynamics and environmental factors. Our findings stress the importance of considering how environmental factors effect species interactions in complex communities. Studies that assess the effect of food web topology on extinction trajectories (e.g. (Dunne & Williams 2009)) can be improved by relating the strength and the sign of species interactions to environmental factors such as temperature (e.g. (Sentis, Hemptinne & Brodeur 2014)). This result underlines the importance of the interplay between indirect interactions and environmental factors on determining cascading extinctions in food webs (Berlow 1999; Dunne & Williams 2009; Fowler 2010).

The extinction time of a species is normally shorter for larger species (Gaston & Blackburn 1995). However, in this experiment we found no significant correlation between species' body size and its predicted mean time to extinction. Though a better model (e.g. an exponential dependence of mean extinction time on species body mass) could be used and these effects further investigated, we find that environmental factors are the main drivers of extinction dynamics in this experiment. The extinction time distribution of the different species could be further correlated to other functional traits of the species and eventually related to measures of functional diversity (Petchey & Gaston 2002). In this way, the combined effect of all relevant traits can be used to characterize the extinction process.

We were able to show that environmental factors were more determinant

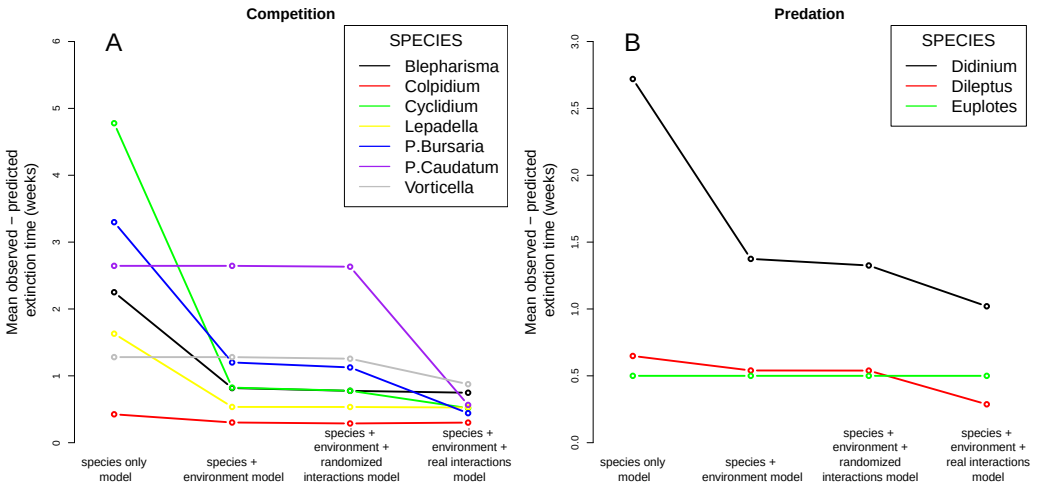

Figure 4.6: Difference in predicted an observed mean time to extinction using "species only model", "species + environment model", "species + environment + randomized interactions model" and "species + environment + real interactions model", for competitors (panel A) and predators (panel B).

than interspecific interactions in determining the extinction risk of the species (figure 4.6). This result underlines the importance of the interplay between indirect interactions and environmental factors on determining cascading extinctions in food webs (Berlow 1999; Dunne & Williams 2009; Fowler 2010). Current research on cascading extinctions has largely considered the topology of food webs as the main driver of extinction cascades, showing the importance of highly connected species (hubs) in determining cascading events Dunne, Williams & Martinez (2002); Allesina & Bodini (2004); Dunne & Williams (2009); Sahasrabudhe & Motter (2011). However, from the results of survival analysis, we suggest that topological effects on cascading extinctions can be investigated together with their interactions with environmental factors such as temperature and nutrient availability (Woodward *et al.* 2010; Sentis, Hemptinne & Brodeur 2014).

The phenomenological model we used in this study (i.e. survival analysis) provided no mechanistic understanding of the extinction process in the food

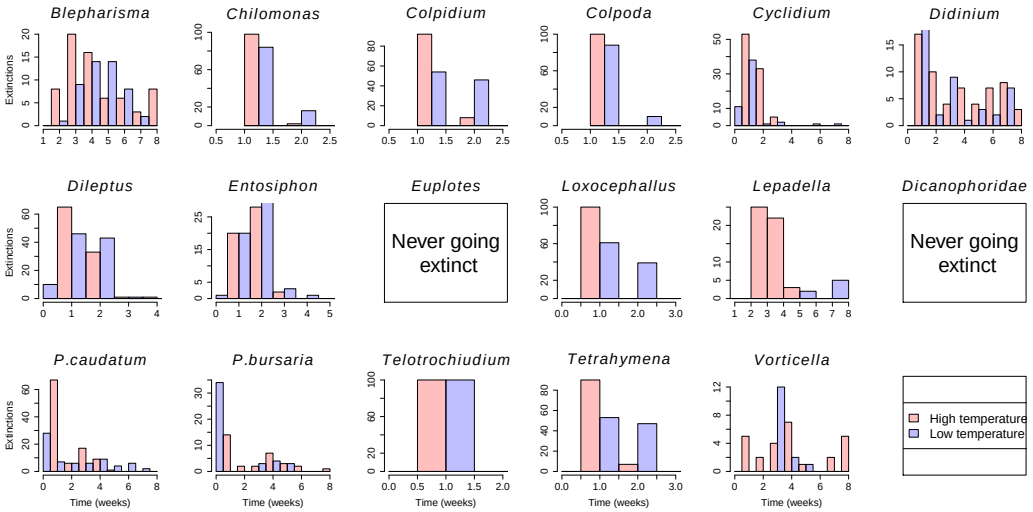

Figure 4.7: Extinction times distributions of the 17 species of the food web. The two colors represent the high (red) and low (blue) temperature treatments

web. Survival analysis gives only a phenomenological understanding of the extinction process and could provide significant insights. This simple model can be further related to more mechanistic models of population dynamics. For example, multispecies stochastic models of population dynamics (e.g. Powell & Roland (2009)) would enable us to relate demographic and environmental parameters to the empirical extinction time distributions and to the topology of the food web. Stochastic models of population dynamics have been proposed for single species (Nåsell 2001; McKane, & Newman 2004, 2005) and multispecies population dynamics (Ebenman, Law & Borrvall 2004; Powell & Roland 2009) and can give fundamentally different results from deterministic models. In the future, we aim to build a generalization of such models (Black & McKane 2012) and to use them to further investigate the data described in this paper.

Bibliography

Allesina, S. & Bodini, A. (2004) Who dominates whom in the ecosystem? Energy flow, bottlenecks and cascading extinctions. *Journal of Theoretical Biology*, **230**, 351-358.

Balvanera, P., Pfisterer, A. B., Buchmann, N., He, J. S., Nakashizuka, T., Raffaelli, D. & Schmid, B. (2006) Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecology Letters*, **9**, 1146-1156.

Belovsky, G. E. (1999) Experimental Studies of Extinction Dynamics. *Science*, **286**, 1175-1177.

Berlow, E. L. (1999) Strong effects of weak interactions in ecological communities. *Nature*, **398**, 330-334.

Black, A. J. & McKane, A. J. (2012) Stochastic formulation of ecological models and their applications. *Trends in Ecology & Evolution.*

Bolker, B. (2013) *Bbmle: Tools for general maximum likelihood estimation.* R package version 1.0-13

Cardillo, M., Mace, G. M., Jones, K. E., Bielby, J., Bininda-Emonds, O. R. P., Sechrest, W., Orme, C. D. L. & Purvis, A. (2005). Multiple causes of high extinction risk in large mammal species. *Science*, **309**, 1239-1241.

Cardinale, B. J., Srivastava, D. S., Duffy, J. E., Wright, J. P., Downing, A. L., Sankaran, M. & Jouseau, C. (2006) Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature*, **443**, 989-992.

Cardinale B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, Venail, P., Narwani, A., Mace, G. M., Tilman, D., Wardle, D. A., Kinzing, A. P., Daily, G. C., Loreau, M., Grace, J. B., Larigauderie, A., Srivastava, D. S. & Naeem, S. (2012) Biodiversity loss and its impact on humanity. *Nature*, **486**, 59-67.

Chapin, F. S., Zavaleta, E. S., Eviner, V. T., Naylor, R. L., Vitousek, P. M., Reynolds, H. L., Hooper, D. U., Lavorel, S., Sala, O. E., Hobbie, S. E., Mack, M. C. & Díaz, S. (2000) Consequences of changing biodiversity. *Nature*, **405**, 234-242.

Daily, G. (1997) *Nature's services. Societal dependence on natural ecosystems.* Island Press, Washington D. C.

Drake, J. M. & Lodge, D. M. (2004) Effects of environmental variation on extinction and establishment. *Ecology Letters*, **7**, 26-30.

Dunne, J. A., Williams, R. J. & Martinez, N. D. (2002) Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology Letters*, **5**, 558-567.

Dunne, J. A. & Williams, R. J. (2009) Cascading extinctions and community collapse in model food webs. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **364**, 1711-1723.

Ebenman, B. & Jonsson, T. (2005) Using community viability analysis to identify fragile systems and keystone species. *Trends in Ecology & Evolution*, **20**, 568-575.

Ebenman, B., Law, R. & Borrvall, C. (2004) Community Viability Analysis: the Response of Ecological Communities To Species Loss. *Ecology*, **85**, 2591-2600.

Fowler, M. S. (2010) Extinction cascades and the distribution of species interactions. *Oikos*, **119**, 864-873.

Gaston, K. J. & Blackburn, T. M. (1995) Birds, body size and the threat of extinction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **347**, 205-212.

Gross, K. & Cardinale, B. J. (2005) The functional consequences of random vs. ordered species extinctions. *Ecology Letters*, **8**, 409-418.

Hanski, I. (1998) Metapopulation dynamics: effects of habitat quality and landscape structure. *Nature*, **79**, 2503-2515.

Hougaard, P., (2000) *Analysis of multivariate survival data.* New York.

Ingram, T. & Steel, M. (2010) Modelling the unpredictability of future biodiversity in ecological networks. *Journal of Theoretical Biology*, **264**, 1047-1056.

Kaunzinger, C. M. K. & Morin, P. J. (2008) Productivity controls food- chain properties in microbial communities. *Nature*, **395**, 495-497.

Kleinbaum D. G. & Klein M., (2005) *Survival Analysis A Self-Learning Text.* Springer Series on Statistics for Biology and Health, New York.

Koh, L. P., Dunn, R. P., Sodhi, N. S., Colwell, R. K., Proctor, H. C. & Smith, V. S. (2004) Species coextinctions and the biodiversity crisis. *Science*, **305**, 1632-1634.

Lande, R. (1993) Risks of Population Extinction from Demographic and Environmental Stochasticity and Random Catastrophes. *The American Naturalist*, **142**, 911-927.

Larsen, T. H., Williams, R. J. & Kremen, C. (2005) Extinction order and altered community structure rapidly disrupt ecosystem functioning. *Ecology Letters*, **8**, 538-547.

Mace, G. M. & Lande, R. (1991) Assessing extinction rates: toward a reevaluation of IUCN treatened species categories. *Conservation Biology*, **5**, 148-157.

McKane, A. J. & Newman, T. J. (2004) Stochastic models in population biology and their deterministic analogs. *Physical Review E*, **70**, 1-19.

McKane, A. J. & Newman, T. J. (2005) Predator-Prey Cycles from Resonant Amplification of Demographic Stochasticity. *Physical Review Letters*, **94**, 1-4.

Millennium Ecosystem Assessment, (2005) *Ecosystem and human well being.* World Resources Institute.

Nåsell, I. (2001) Extinction and quasi-stationarity in the Verhulst logistic model. *Journal of Theoretical Biology*, **211**, 11-27.

Newman, M. E. J. (1996) Self-organized criticality, evolution and the fossil extinction record. *ArXiv*, **3801**, 1-20.

Newman, M. E. J. (2000) Random graphs as models of networks. ArXiv, **1**, 1-35.

Paine, R. T. (1966) Food Web Complexity and Species Diversity. *The American Naturalist*, **100**, 65.

Petchey, O. L. & Gaston, K. J. (2002) Extinction and the loss of functional diversity. *Proceedings of the Royal Society B: Biological Sciences*, **269**, 1721-1727.

Petchey, O. L., Gonzalez, A. & Wilson, H. B. (1997) Effects on population persistence: the interaction between environmental noise colour, intraspecific competition and space. *Proceedings of the Royal Society B: Biological Sciences*, **264**, 1841-1847.

Pimm, S. L., Lones, H. L. & Diamond, J. (1988) On the risk of extinction. *The American Naturalist*, **132**, 757-785.

Pimm, S. L., Jenkins, C. N., Abell, R., Brooks, T. M., Gittleman, J. L., Joppa, L. N., Raven, P. H., Roberts, C. M. & Sexton, J. O. (2014) The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, **344**,

Powell, C. R. & Roland, R. P. (2009) The effects of stochastic population dynamics on food web structure. *Journal of Theoretical Biology*, **257**, 170-180.

Purvis, A., Gittleman, J. L., Cowlishaw, G. & Mace, G. M. (2000) Predicting extinction risk in declining species. *Proceedings of the Royal Society B: Biological Sciences*, **267**, 1947-1952.

Raup, D. M. & Sepkoski, J. J. (1984) Periodicity of extinctions in the geologic past. *Proceedings of the National Academy of Sciences of the United States of America*, **81**, 801-805.

Rybicki, J. & Hanski, I., (2013) Species-area relationships and extinctions caused by habitat loss and fragmentation. *Ecology Letters*, **16** 27-38.

Sabo, J. L. (2008) Population viability and species interactions: Life outside the single-species vacuum. *Biological Conservation*, **141**, 276-286.

Sahasrabudhe, S. & Motter, A. E. (2011) Rescuing ecosystems from extinction cascades through compensatory perturbations. *Nature Communications*, **170**, 1-21.

Sentis, A., Hemptinne, J. L. & Brodeur, J. (2014) Towards a mechanistic understanding of temperature and enrichment effects on species interaction strength, omnivory and food-web structure. *Ecology Letters*, **17**, 785-793.

Solan, M., Cardinale, B. J., Downing, A. L., Engelhardt, K. A. M., Ruesink, J. L. & Srivastava, D. S. (2004) Extinction and ecosystem function in the marine benthos. *Science*, **306**, 1177-1180.

Srinivasan, U. T., Dunne, J. A., Harte, J. & Martinez, N. D. (2007) Response of complex food webs to realistic extinction sequences. *Ecology*, **88**, 671-682.

Thompson, A S., Rhodes, J. C. & Pettman, I. (1988) Culture collection of algae and protozoa; a catalogie of strains (5th editio) CCAP, *Freshwater Biology*.

Worsfold N., Ph.D. Thesis (2007) *The consequences of extinctions in experimental aquatic communities.*

Woodward, G., Benstead, J. P., Beveridge, O. S., Blanchard, J., Brey, T., Brown, L. E. & Yvon-durocher, G. (2010) Ecological networks in a changing climate. *Advances in Ecological Research*, **42** 72-122.

Zavaleta, E. S. & Hulvey, K. B. (2004) Realistic species losses disproportionately reduce grassland resistance to biological invaders. *Science*, **306**, 1175-1177.

General Conclusion

In this thesis we presented a collection of studies where stochastic methods, mostly the theory of Birth and Death processes, have been applied to describe population dynamics and to infer relevant ecological parameters from population time series data. We investigated fundamental ecological questions related to environmental dependencies, demographic stochasticity, invasion and extinction dynamics using several modifications of the continuous time stochastic logistic birth and death process (Nåsell 2001).

In the first chapter we used a temperature dependent version of the stochastic logistic process showing how experimental factors and inference methods interact to influence the accuracy with which activation energy, the main driver of metabolic responses to temperature (Brown *et al.* 2004), can be inferred. We found that the fraction of habitat searched is the most important factor in determining the accuracy of the estimates of activation energy. We also developed a novel method to infer activation energy that we named the "direct method" that is inferring directly the activation energy from different population time series at different temperatures. Our findings are likely to be a step forward in developing a robust predictive framework for predicting the response of populations to climate change. The methods developed in the first chapter can be further tested with different processes such as feeding rates (Rall *et al.* 2009; Englund *et al.* 2011; Fussmann *et al.* 2014) and with different environmental variables such as nutrient concentration (Weisse *et al.* 2002). Eventually, our approach can be used in more complex models of population and community dynamics taking into account the real determinants of climate change (Pereira *et al.* 2010).

In the second chapter, we added interspecific competition to the birth rate of the stochastic logistic process providing an explanation of how demographic stochasticity biases our predictive understanding of invasion dynamics. Moreover, we presented a novel method to infer competition parameters from the first stages of an invasion, taking into account the effect of demographic stochasticity. The method we developed can be used to infer competition parameters when one of the competing species remains at its carrying capacity, and the other species is at low density. The approximations we developed are particularly useful in modeling other factors that might affect invasion dynamics resulting from competition (Simberloff *et al.* 2013; Lurgi *et al.* 2014), such as density dependent effects, Allee effects (Drake 2004; Drake & Lodge 2006) and environmental stochasticity (Drake *et al.* 2006). Proper microcosm experiments can be developed from this study in order to investigate such effects (Drake *et al.* 2011). Moreover, combined with the methods developed in the first chapter, our method can be used to further investigate the temperature dependence of competitive interactions (Fox & Morin 2001).

In the third chapter we added the effects of trophic interaction to the death rate of the stochastic logistic process, showing how different choices of the foraging parameters vary the mean time to extinction by up to 10 orders of magnitude, even when equilibrium population size is kept constant. Our results imply that estimates of extinction risk could be extremely inaccurate without explicitly accounting for trophic interaction and may have critical implications in conservation planning (Sabo & Gerber 2007; Sabo 2008). From a mathematical perspective, more accurate methods derived from standard techniques used in perturbative quantum mechanics (WBK or eikonal approximations) (Assaf & Meerson 2007) can be used to describe extinction dynamics. A broad class of models of single-species populations in a static environment has been systematically analysed yielding analytical expressions for the MTE (Assaf & Meerson 2010) and has been extended to interacting populations (Khasin & Dykman 2009; Parker & Kamenev 2009). The problem of describing the dynamics of interacting stochastic populations is a

hot topic in physical research (Khasin & Dykman 2009) and the analytical solutions presented in the third chapter for the simplest cases can give specific insights in the results obtained with more refined techniques (Parker & Kamenev 2009). For example the model we used to describe predator prey interaction, when using a type III functional response, presents interesting bistable dynamics that have been further investigated by Bruna, Cahpman & Smith (2014).

In the last chapter of the thesis, we presented an experiment with a small microbial food web, where both competition and predation affect the extinction rate distribution of all the species. Given the complexity of the system, we used a phenomenological model that takes into account both environmental factors and species interactions, but not population dynamics. From the phenomenological analysis, we found that environmental factors are more important than species' interactions in determining the extinction times distribution of the species. We aim to extend the models derived in the first three chapters to describe multispecies population dynamics and eventually the extinction times distribution observed in the experiment described in the last chapter. We aim to derive a set of multispecies stochastic models that take into account the different effects of ecological interactions. As an example of application, such a set of stochastic models will enable us to construct temperature dependent biodiversity scenarios for ecological communities. The corresponding inference scheme will provide an assessment of the predictive understanding we can gather from the application of such models. We give a particular focus on the importance of using microcosms experiments as biological models. Microcosms experiments have played a fundamental role in shaping ecological theory, providing elegant explanation of the mechanisms shaping ecological communities (Drake, Huxel & Hewitt 1996; Benton *et al.* 2007; Drake & Kramer 2011). Such experiments have been used to disentangle the mechanisms behind extinctions (Holyoak *et al.* 2000; Bell & Gonzalez 2009; Drake *et al.* 2011; Ferguson & Ponciano 2014) and to test extinction predictors (Clements *et al.* 2013). As new complex models of population and community dynamics are produced, proper micro-

cosms experiments can be designed to check new models' assumptions and predictions.

We used single species stochastic models as they can be handled analytically (McKane, & Newman 2004; Black & McKane 2012). The use of stochastic models such as continuous birth and death processes can provide fundamentally different results from deterministic models (Ebenman, Law & Borrvall 2004; McKane, & Newman 2005). Moreover, single species models provide a probabilistic framework to derive inference schemes from (Ross, Taimre & Pollett 2006; Ross, Pagendam & Pollett 2009) still providing useful insight into the determinants of population dynamics (Black & McKane 2012). Using approximations, such as the diffusion one, we provided an analytical expression for the first two moments of the population probability distribution of interacting species (Ross, Pagendam & Pollett 2009; Ross 2012) and thus we were able to have a reliable, but still approximated, analytic expression for the likelihood function of the models. The use of continuous time stochastic models has been proposed for many other biological systems, including evolutionary processes (Bartlett 1949; Kimura 1964) and epidemics (Bailey 1957; Anderson & May 1992). However, Their application in ecology is limited partly due to a lack of clear statistical procedures for fitting those models to data (Ross, Taimre & Pollett 2006). In this thesis, we provide novel and robust statistical methods to fit such models to population time series data. The inference scheme developed give a complete estimation of the probability distribution of the parameters of the models, including interaction parameters when they are present, thus enabling the assessment of the predictive power of the models given the available data.

4.6 Future research

. One of the biggest frontiers of ecology is being able to predict the properties and dynamics of ecosystems and their components under perturbations (Beckage, Gross & Kauffman 2011). To do that we need to understand the ef-

fects of demographic and environmental stochasticity, of species interactions and of how we obtain and formulate our models in order to take into account all these factors. In this thesis, we put the foundations for a predictive understanding of the relative effect of such factors. Using the simple models developed in this thesis together with the associated inference methods, we aim to end up with a theoretical framework in which one could predict the probability of extinction or other events within complex food webs.

The steps needed to achieve this goal are summarized in the following three points:

- Building stochastic models of multispecies population dynamics in complex food webs. Such models need to incorporate stochastic population dynamics, species' interactions, and environmental variation and can be formulated in the form of birth and death processes. Particular care should be given to the parameterization of such models. For example, in order to reduce the high number of parameters that naturally arises when many species are present, relevant scaling factors across species can be used (e.g. allometric scaling relations (Hudson & Reuman 2013)). The topology of the network of interactions can then be investigated in relation to the possible population dynamics (Ebenman, Law & Borrvall 2004). Approximations can be used when the number of individuals of a given species is sufficiently high, so that the stochastic effects can be neglected (Black & McKane 2012).

- Building inference frameworks associated to the stochastic models. Once the proper model describing multispecies population dynamics is obtained, the likelihood function of the model can be simulated using the Gillespie algorithm (Gillespie 1977) and time series data obtained from model simulations can be used to test our ability to infer population parameters and predict community dynamics. In the case of complex multispecies models we cannot obtain the likelihood function in a closed form, as we did for single species models, but we can use more refined, but also more computationally demanding, inference techniques such

Approximate Bayesian Computation (Beaumont 2010) to simulate the likelihood. Predictions of the extinction times distribution, like the one described in the last chapter of the thesis, can be explicitly formulated using such inference frameworks.

- Testing model predictions with proper microcosm experiments. From the previous steps, we can gather the information about the quality and quantity of data that is needed to make reliable predictions about a specific mechanism driving the dynamics of an ecological community. Then we can design a corresponding experiment to test the model's predictions against real data. Microcosm experiments are particularly useful to achieve this goal as they can be manipulated in both species composition and environmental factors (Drake & Kramer 2011).

Building such a multispecies stochastic model of population dynamics, describing the effect of temperature in multispecies communities of competitors, will be the subject of my first Postdoc. The project is entitled "Improving ecological predictions via enhanced inference methods" and has been selected by the Forschungskredit Committee of the University of Zürich.

Bibliography

Anderson, R. M. & May R. M. (1992) *Infectious diseases of humans: dynamics and control*. Oxford University Press.

Assaf, M. & Meerson, B. (2007) Spectral theory of metastability and extinction in a branching-annihilation reaction. *Physical Review E*, **75**, 1-8.

Assaf, M. & Meerson B. (2010) Extinction of metastable stochastic populations. *Physical Review E*, **81**, 1-18.

Bailey, N. T. J. (1957) *The mathematical theory of epidemics*. Griffin, London.

Bartlett, M. S. (1949) Some evolutionary stochastic processes. *Journal of the Royal Statistical Society, Series B (Methodological)*, **11**, 211-229.

Beaumont, M. A. (2010) Approximate Bayesian Computation in evolution and ecology. *Annual Review of Ecology, Evolution, and Systematics*, **41**, 379-406.

Beckage, B., Gross, L. J. & Kauffman, S. (2011) The limits to prediction in ecological systems. *Ecosphere*, **2**, 1-12.

Bell, G. & Gonzalez, A. (2009). Evolutionary rescue can prevent extinction following environmental change. *Ecology Letters*, **12**, 942-948.

Benton, T. G., Solan, M., Travis, J. M. J. & Sait, S. M. (2007) Microcosm experiments can inform global ecological problems. *Trends in Ecology & Evolution*, **22**.

Black, A. J. & McKane, A. J. (2012) Stochastic formulation of ecological models and their applications. *Trends in Ecology & Evolution*, **27**, 337-345.

Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M. & West, G. B. (2004) Toward a Metabolic Theory of Ecology. *Ecology*, **85**, 1771-1789.

Bruna, M., Chapman, S. J. & Smith, M. J. (2013) Model reduction for slow-fast stochastic systems with metastable behaviour. *The Journal of Chemical Physics*, **140**, 174107.

Clements, C. F., Worsfold, N. T., Warren, P. H., Collen, B., Clark, N., Blackburn, T. M. & Butler, S. (2013) Experimentally testing the accuracy of an extinction estimator: Solow's optimal linear estimation model. *The Journal of Animal Ecology*, **82**, 345-354.

Drake, J. M. & Kramer, A. M. (2011) Mechanistic analogy: how microcosms explain nature. *Theoretical Ecology*, **5**, 433-444.

Drake, J. M. & Kramer, A. M. (2011) Mechanistic analogy: how microcosms explain nature. *Theoretical Ecology*, **5**, 433-444.

Drake, J. M. & Lodge, D. M. (2006) Allee effects, propagule pressure and the probability of establishment: risk analysis for biological invasions. *Biological Invasions*, **8**, 365-375.

Drake, J. M., Drury, K. L. S., Lodge, D. M., Blukacz, A., Yan, N. D. & Dwyer, G. (2006) Demographic stochasticity , environmental variability , and windows of invasion risk for Bythotrephes longimanus in North America. *Biological Invasions*, **8**, 843-861.

Drake, J. M., Shapiro, J. & Griffen, B. D. (2011) Experimental demonstration of a two-phase population extinction hazard. *Journal of the Royal Society, Interface*.

Drake, J. M. (2004) Allee effects and the risk of biological invasion. *Risk Analysis*, **24**, 795-802.

Drake, J. A., Huxel, G. R. & Hewitt, C. L. (1996) Microcosms as Models for Generating and Testing Community Theory. *Ecology*, **77**, 670-677.

Ebenman, B., Law, R. & Borrvall, C. (2004) Community viability analysis: the response of ecological communities to species loss. *Ecology*, **85**, 2591-2600.

Englund, G., Ohlund, G., Hein, C. L. & Diehl, S. (2011) Temperature dependence of the functional response. *Ecology Letters*, **14**, 914-921.

Ferguson, J. M. & Ponciano, J. M. (2014) Predicting the process of extinction in experimental microcosms and accounting for interspecific interactions in single-species time series. *Ecology Letters*, **17**, 251-259.

Fox, J. W. & Morin, P. J. (2001) Effects of intra- and interspecific interactions on species responses to environmental change. *Journal of Animal Ecology*, **70**, 80-90.

Fussmann, K. E., Schwarzmüller, F., Brose, U., Jousset, A. & Rall, B. C. (2014) Ecological stability in response to warming. *Nature Climate Change*, **4**, 206-210.

Holyoak, M., Lawler, S. P. & Crowley, P. H. (2000) Predicting Extinction: Progress With an Individual-Based Model of Protozoan Predators and Prey. *Ecology*, **81**, 3312-3329.

Hudson, L. N. & Reuman, D. C. (2013) A cure for the plague of parameters: constraining models of complex population dynamics with allometries. *Proceedings of the Royal Society Series B: Biological Sciences*, **280**, 1-8.

Khasin M. & Dykman M. I. (2009) Extinction rate fragility in population dynamics. *Physical Review Letters*, **103**, 1-4.

Kimura, M. (1964) Diffusion models in population genetics. *Journal of Applied Probability*, **2**, 177-232.

Lurgi, M., Galiana, N., López, B. C., Joppa, L. & Montoya, J. M. (2014) Network complexity and species traits mediate the effects of biological invasions on dynamic food webs. *Manuscript*

McKane, A. J. & Newman, T. J. (2004) Stochastic models in population biology and their deterministic analogs. *Physical Review E*, **70**, 1-19.

McKane, A. J. & Newman, T. J. (2005) Predator-prey cycles from resonant amplification of demographic stochasticity. *Physical Review Letters*, **94**, 1-4.

Nåsell, I. (2001) Extinction and quasi-stationarity in the Verhulst logistic model. *Journal of Theoretical Biology*, **211**, 11-27.

Parker, M. & Kamenev, A. (2009) Extinction in the Lotka-Volterra model. *Physical Review E*, **80**, 1-10.

Pereira, H. M., Leadley, P. W., Proença, V., Alkemade, R., Scharlemann, J. P. W., Fernandez-Manjarrés, J. F., Araújo, M. B., Balvanera, P., Biggs, R., Cheung, W. W. L., Chini, L., Cooper, H. D., Gilman, E. L., Guénette, S., Hurtt, G. C., Huntington, H. P., Mace, G. M., Oberdoff, T., Revenga, C., Rodrigues, P., Scholes, R. J., Rashid Sumalia, U. & Walpole, M. (2010) Scenarios for global biodiversity in the 21st century. *Science*, **330**, 1496-1501.

Rall, B. C., Vucic-Pestic, O., Ehnes, R. B., Emmerson, M. & Brose, U. (2009) Temperature, predator-prey interaction strength and population stability. *Global Change Biology*, **16**, 2145-2157.

Ross, J. V. (2012) On parameter estimation in population models III: time-inhomogeneous processes and observation error. *Theoretical Population Biology*, **82**, 1-17.

Ross, J. V., Pagendam, D. E. & Pollett, P. K. (2009) On parameter estimation in population models II: multi-dimensional processes and transient dynamics. *Theoretical Population Biology*, **75**, 123-132.

Ross, J. V., Taimre, T. & Pollett, P. K. (2006) On parameter estimation in population models. *Theoretical Population Biology*, **70**, 498-510.

Sabo, J. L., & Gerber, L. R. (2007) Predicting extinction risk in spite of predator-prey oscillations." Ecological Applications **17**, 1543-1554.

Sabo, J. L. (2008) Population viability and species interactions: Life outside the single-species vacuum.

Simberloff, D., Martin, J. L., Genovesi, P., Maris, V., Wardle, D. A., Aronson, J., Courchamp, F., Galil, B., García-Berthou, E., Pascal, M., Pyšek, P., Sousa, R., Tabacchi, E. & Vilá, M. (2013) Impacts of biological invasions: what's what and the way forward. *Trends in Ecology & Evolution*, **28**), 58-66.

Weisse, T., Stadler, P., Lindström, E. S., Kimmance, S. A. & Montagnes, D. J. S. (2002) Interactive effect of temperature and food concentration on growth rate: A test case using the small freshwater ciliate Urotricha farcta. *Limnology and Oceanography*, **47**, 1447-1455.

Acknowledgements

It is customary to start with academic thanks; for this I start with Owen Petchey, without his support this thesis would have never started. Owen taught me that the most important quality of a good scientist is the ability to communicate effectively his science with others. Coming from theoretical physics and working in a group of empirical ecologists I needed to develop my communication skills. In the beginning it was really hard to find a common language between physics and ecology. I am grateful to Owen and to the other members of the group for their patience and for their continuous will to discuss despite my ignorance in biology. After many discussions where we didn't understand with each other, we managed to work together in a positive way, using our different backgrounds for a common scientific goal. Being Owen's first Ph.D. student in Zürich has been another challenge for me; I saw the group growing, new people arriving and people leaving.

My Ph.D. was generously funded by Microsoft Research and I spent some time working directly at the computational ecology group in Cambridge. I always felt tremendously stimulated from the work of people there. A special thank, therefore, goes to Matthew Smith, my supervisor at Microsoft Research. Matthew, with his constant presence and willingness to discuss everything, encouraged and supported me, passing me his enthusiasm for ecology. I take this occasion to thank also Gustav Delius, who assisted me in the first stages of my transition from physics to ecology. I have good memories of our meetings at the mathematics department of the University of York and in Berlin.

I want to thank also all the new friends and colleagues that I met during these years in Zürich, Sheffield, Cambridge and all over the world. thanks to Marco Plebani who shared with me office, joys and sorrows, thanks to Chris Clements, Nick Worsfold, Emily Griffiths, Aaron Thierry, Christine Thuring, thanks to Enrica De Luca, Piero Visconti, Antonio Scala, Aabir Banerji, Dennis Hansen ... and many many others, the list is too long to be put in this page. I have been lucky to have people around me that supported me during the bad moments and celebrated with me the good moments.

Un ultimo ringraziamento vá alla mia famiglia, a mia madre, mio padre e mio fratello, ed ai miei amici di Roma, a Giuseppe e a Viviana, che hanno sopportato tutte le mie assenze, i miei malumori e le mie lune storte. Durante questi anni ho trovato sempre un porto sicuro in voi. E per concludere ringrazio Beatrice, a cui questa tesi é dedicata. Senza il tuo sostegno, continuo e costante, non sarei mai arrivato fin qui. Grazie di cuore.

Curriculum Vitae

Gian Marco PALAMARA

Born in Rome the 27-11-1980, Italian
Current address:
Weinbergstrasse, 84
8006 Zürich
Switzerland
Mobile: +41 (0) 78 677 51 17
Email: gianmarco.palamara@ieu.uzh.ch

Education

2011 - present	**Ph.D. Candidate, Integrative Ecology** Department of Evolutionary Biology and Environmental Studies (IEU), University of Zürich, Switzerland
2010 - 2011	**PhD Candidate, Ecology** Department of Animal and Plant Sciences (APS), University of Sheffield, United Kingdom
2006 - 2009	**Master Degree, Theoretical Physics** University of Rome, "La Sapienza", Rome, Italy Graduated 110/110
2003 - 2006	**Bachelor Degree, Physics** University of Rome, "La Sapienza", Rome, Italy Graduated 105/110
2002 - 2003	**Erasmus student, Fluid Mechanics** Université Paul Sabatier, Toulouse, France
2000 - 2002	**Undergraduate student, Mechanical Engineering** Engineering Faculty, University of Rome, "La Sapienza", Rome, Italy
1999	**Maturitá Scientifica** Collegio San Giuseppe, Istituto De Merode, Rome, Italy

Professional Experience

2011 - present	**PhD Student, University of Zürich, Switzerland.** Works at the IEU, Developing computational Models to describe the collapse of ecological communities. Teaching assistant for ecology block courses and programming courses

September-December 2013	**Research Intern, Microsoft Research, Cambridge, UK** Worked at the Computational Ecology and Environmental Sciences (CEES) Group, developing software for the statistical inference of ecological parameters in population models.
March-July 2010	**Intern, National Commission for UNESCO, Rome, Italy** Event planning for the International Year of Biodiversity
2008 - 2010	**Research Assistant, University of Rome "La Sapienza", Italy** Worked in the PIL group at the Department of Physics, developing numerical models of population dynamics and food web structure.

Publications

Palamara, G. M., Childs, D. Z., Clements, C.F., Petchey, O. L., Plebani, M. & Smith, M. J. Inferring the temperature dependence of ecological parameters: the effect of experimental design and inference algorithm. *Ecology &Evolution*, In Press.

Palamara, G. M., Delius, G. W., Smith, M. J. & Petchey, O. L. (2013) Predation effects on mean time to extinction under demographic stochasticity. *Journal of Theoretical Biology*, **334**, 61-70.

Palamara, G. M., Zlàtic, V., Scala, A. & Caldarelli, G. (2011) Population dynamics on complex food webs. *Advances in Complex Systems*, **14**, 635-647.

Awards and Fellowships

2015	Postdoc Grant. Improving ecological predictions via enhanced inference methods. Forschungskredit, University of Zürich
2010 - 2013	Microsoft Research Ph.D. Fellowship
September 2010	Best Paper Award, ECSS' 10, 2nd Prize

Scientific Presentations at International Conferences

September 2014	Oral presentation.Extinction rates in an experimental microbial aquatic comunity. **European Conference on Complex Systems 2014**, Lucca, Italy
August 2013	Oral presentation. Extinction rates in an experimental microbial aquatic community. **INTECOL 2013**, London, UK
August 2013	Oral presentation.Estimating the environmental dependence of ecological parameters. **INTECOL 2013**, London, UK

December 2012	Oral presentation. Estimating the environmental dependence of ecological parameters. **BES Annual Meeting 2012**, Birmingham, UK
August 2012	Oral presentation. Predation effects on mean time to extinction under demographic stochasticity. **ESA Annual Meeting 2012**, Portland, Oregon, USA
September 2011	Oral presentation. Understanding the effect of functional response on the mean time to extinction**.** **BES Annual Meeting 2011**, Sheffield, UK
August 2011	Poster presentation. Effect of functional Response to mean time to extinction. **ESA Annual Meeting 2011**, Austin, Texas, USA
September 2010	Oral presentation. Population dynamics on complex food webs. **European Conference on Complex Systems 2010**, Lisbon, Portugal

Workshops

8-10 September 2014	**ECO**2 Symposium: Ecology and Economics.** London School of Economics, London, UK
07-18 July 2014	**Ramon Margalef Summer Colloquia.** Institute of Marine Sciences, Barcelona, Spain.
13-16 May 2013	**Workshop: Turining Data into Insights.** Microsoft Research & FAPESP, São Paulo, Brazil
9-12 April 2013	**2nd UNCCD Scientific Conference.** Bonn, Germany. Worked as Rapporteur for the Global Risk Forum of Davos
3-9 February 2013	**Winter School: Ecological Theory.** Annual retreat of the PhD Program in Ecology, Monte Veritá, Switzerland
3-7 September 2012	**Summer School: Stochastic Modelling for Systems Biology**. Applied Bayesian Statistics School (ABS), Pavia, Italy
21-25 May 2012	**Workshop: Stochastic Dynamics in Action**. Zentrum für interdisziplinäre Forschung (ZIF), Bielefeld, Germany
15 March 2012	**2nd Protist Workshop**, EAWAG Aquatic Research, Zürich, Switzerland
27-30 June 2011	**Microsoft Research Summer School**, Microsoft Research, Cambridge, UK

6-10 June 2011	**International School and Conference on Network Science**, (NETSCI) Budapest, Hungary.
9-15 April 2011	**SIZEMIC Workshop and Summer School**, Institute of Hydrobiology and Fisheries Science, Hamburg, Germany

Visited Institutes

April 2014	**Ralph M. Parsons Laboratory, Massachusetts Institute of Technology (MIT)**, Boston, USA
April 2013	**Laboratory of Eco-Hydrology, EPFL**, Lausanne, Switzerland
March 2013	**Theoretical Evolutionary Ecology Group, EAWAG.** Kastanienbaum, Switzerland
July 2011	**Food and Agriculture Organization of the United Nations (FAO)**, Rome, Italy
June 2011	**Ruder Boščović Institute, Theoretical Physics Division**, Zagreb, Croatia
October 2010 - June 2011	**Mathematics Department, University of York**, York, UK
March 2010	**Department of Ecology and Ecosystem Modelling, University of Potsdam,** Potsdam, Germany
August 2009	**Istituto Nacional de Pesquisa Amazonica (INPA)**, Manaus, Brasil

Additional Relevant Competences

Languages	Italian (native), French (native), English (fluent), Spanish (intermediate), Portuguese (Intermediate), Catalan (intermediate), German (Beginner)
Programming	C, C++, C#, R, Matlab/Octave, Mathematica, LaTeX
Operating systems	Linux, Windows (Microsoft Office)
Simulations	Monte Carlo, Stochastic Dynamics, Differential Equations, Numerical Methods, Graphs and Networks
Other skills	Interrelation in multi-cultural environment. Suited for team working